Cell Type Diversity During Sea Urchin Development: A Single Cell Approach to Reveal Different Neuronal Types and Their Evolution

Thesis

The Open University
School of Life, Health and Chemical Sciences
Milton Keynes, United Kingdom
Affiliated Research Center:
Stazione Zoologica Anton Dohrn
Naples, Italy

Cell Type Diversity During Sea Urchin Development: A Single Cell Approach to Reveal Different Neuronal Types and Their Evolution

Periklis Paganos, MSc

Doctor of Philosophy

June 2021

This PhD project has been carried out at Stazione Zoologica Anton Dohrn under Dr. Maria Ina Arnone's supervision.

Supervisory team

Director of Studies: Dr. Maria Ina Arnone, Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Naples, Italy

External Supervisor: Prof. Gáspár Jékely, Living Systems Institute, University of Exeter, Exeter, United Kingdom

Abstract

An essential step towards understanding the biology and development of an organism is identifying the molecular fingerprint of its cell types, while to understand its evolution, cell type inventories must be obtained and compared across taxa. This thesis aims to provide a thorough characterization of the genetic program employed in the cell types of a non-chordate deuterostome, the sea urchin. To this end, single cell RNA sequencing (scRNA-seq) was used as a tool to recognize the major cell types in place during the late embryonic and early larval development of the sea urchin *Strongylocentrotus purpuratus*. The outcome of this thesis was the generation of scRNA-seq cell type atlases for the late gastrula and two early pluteus larval stages, depicting their specific molecular signatures. Extensive analysis of the scRNA-seq data revealed complex regulatory states and novel gene markers. Moreover, the scRNA-seq analysis showed that known gene interactions could be traced back at a single-cell level and that novel functional domains of known gene regulatory modules can be described. In addition, a thorough analysis of the nervous system showed signs of increased neuronal complexity and diversity and identified pre-neuronal cell types employing a similar genetic program to the one utilized by the vertebrate cells that will give rise to the forebrain region. Furthermore, a detailed analysis of the early larva cell types led to identifying specific gastrointestinal and neuronal larval populations with a pancreatic-like genetic program. Lastly, the data presented in this thesis demonstrate that the sea urchin homolog of the master regulator *Pdx1*, which in vertebrates controls endocrine pancreas differentiation, has an evolutionary conserved role in promoting secretory fate by regulating the differentiation of a specific neuronal type with a strong pancreatic-like molecular signature.

To my family and friends that are always by my side

Acknowledgments

This PhD thesis would not be possible without the support of the following people, whom I would like to thank from the bottom of my heart to support, help, and shape me into the person I am today.

Above all, I would like to thank my mentor and supervisor, Dr. Maria Ina Arnone, to support, guide, and trust me throughout my stay at Stazione Zoologica. Immediately from the beginning, Ina put a tremendous amount of trust in me, let me work independently, and never discouraged me from pursuing even the craziest ideas. Thank you for showing me that a person can be successful and still enthusiastic, that a person can be strong and, at the same time, one of the kindest people with the warmest hearts. Thank you for giving me a once in a lifetime opportunity and providing the perfect environment to pursue my dreams. Thank you for welcoming me to your lab family and your home.

I would also like to express my gratitude to my external supervisor Professor Gáspár Jékely, whose advice and valuable insights had a significant impact on shaping o this project.

I would like to thank European Union's Horizon 2020 research and innovation program for generously funding my studies [Marie Skłodowska-Curie grant agreement no. 766053 (EvoCELL)].

I would like to thank Dr. Detlev Arendt for the fantastic collaboration of the scRNA-seq project and for hosting me in his lab. I would also like to thank Drs. Jake Musser and Francesco Lamanna for their patience while teaching me how to perform computational analysis.

I am grateful to Jovana Ranđelović, Danila Voronov, Maria Cocurullo for the life in and outside of the lab, for their practical and emotional support. Most importantly, I thank them for being "my first international friends", for being real and irreplaceable friends, filling my day with joy and laughter even from afar.

I thank Dr. Margherita Perillo for the work I performed with her and for being such a friendly and supportive person.

I would also like to thank Dr. Claudia Cuomo, who was the first to welcome me to Italy, guide me to the lab, and introduce me to the Neapolitan lifestyle, always with a smile on her face.

I also thank Dr. Giovanna Benvenuto for her advice, sharing her expertise, and overall amazing and one of the most joyful collaborations. I would also like to say thank you to Drs. Salvatore D'Aniello and Filomena Caccavale for the excellent collaboration on the Nitric oxide project and their advice and teaching.

I am also grateful to our extended lab family member Professor Paola Oliveri for our fruitful and stimulating conversations, scientific arguments, advice, collaboration, and significant input in shaping the project. I would also like to thank Dr. Enrico D'Aniello for his advice and fun partnership writing the review article.

I would also express my gratitude to Drs. Vladimir Benes and Bianka Baying at GeneCore, EMBL, and the Molecular Biology Service, as well as the Advanced Microscopy Center members at SZN, for their precious help and contribution to this project. In particular, I would like to thank Pasquale De Luca, Elvira Mauriello, Raimondo Pannone, Elio Biffali, Giovanni Gragnaniello, Giampiero Lanzotti, and Franco Iamunno.

I also thank Mara Francone and Alessandro Amoroso for ensuring that I was stocked with general supplies and Davide Caramiello for taking good care of the adult sea urchins.

I am also grateful to all the EvoCELL ITN members (fellows, PIs, and our project manager Francesca Stomeo) for our stimulating interactions, training activities, and fun we had over the last three years.

I would also like to thank my examiners for their valuable input and recommendations that helped shape the final version of my thesis.

Apart from these people who contributed to this thesis's outcome throughout my three-year stay at Stazione Zoologica, I would also like to take the opportunity and thank some people in my life that made this thesis possible.

I would like to thank with all my heart my first mentor, captain, and very dear friend Dr. Constantin N. Flytzanis, for introducing me to the field of molecular biology and sea urchin development. Costas has played a crucial role in my personal and scientific development by teaching me almost everything I know regarding sea

urchin development and showing me how to pursue my dreams, and make me believe that I am worthy of dreaming. Costas changed the course of my life, and for that, I will be forever grateful.

I would also like to thank my closest friends (probably the term family would be more appropriate) Angeliki Klopa, Marieta Karanika, and Eugenia Vintzilaiou for their support and love over the years. The American novelist S. E. Hinton once said, “If you have two friends in your lifetime, you are lucky. If you have one good friend, you are more than lucky. I believe that he was wrong because having these three people in my life makes me feel like the luckiest person in the world.

Finally, I would like to thank my parents Peggy and Andreas, and my brother Panagiotis for their unconditional love, support, and faith in me. Without you this thesis would not have been possible.

Table of Contents

List of Tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 5

Chapter 1

Introduction

This chapter contains a general introduction of the thesis, introduces fundamental concepts necessary for understanding the aims of this thesis, and presents sea urchin as an experimental model for studying development and evolution. Moreover, it provides information on sea urchin cell type specification and differentiation, gene regulation, and the omics approaches used. This chapter also contains the aims of this thesis.

1.1. Gene expression

The identity and survival capability of all living forms depend on the genetic information (genes) encoded in their genome (DNA) and how this information is converted into functional units (RNAs and proteins). Gene expression is a universal and highly regulated process by which the genetic information flows into effector products. The first step of this process is the transcription of a gene into RNA molecules that can either be coding or non-coding. The main difference is that coding RNAs are an intermediate stage necessary to form functional products (proteins), whereas non-coding RNAs (ncRNAs) are the functional units themselves. Coding RNAs (mRNAs) are used as a substrate to produce peptides or polypeptidergic molecules (proteins) through the translation process.

1.2. Embryonic development

Animal embryonic development, or else known as embryogenesis, is a term used to describe a series of dynamic processes leading to the formation of an organism starting from a single cell. In the case of all sexually reproduced animals, that single cell is the fertilized oocyte. However, despite the common origin of embryogenesis for such animals, morphologically similar oocytes give rise to different organisms. Moreover, within the same organism, the fertilized oocyte gives rise to cell types forming diverse tissues, organs, and body parts. The source of this diversity lies in differential gene expression. All cells within an organism share the same DNA and genes. What differentiates a cell from another is the genes expressed within them, and the differential expression is a result of differential gene regulation. Thus unraveling the mechanisms that regulate gene expression across cells is crucial for understanding their development and evolution.

1.3. Cell type evolution

Multicellularity has evolved as an advantageous feature supporting life, while its evolutionary origin remains a mystery. Several evolutionary theories converge that all life forms originated from unicellular organisms that, in order to survive had to co-operate with other unicellular organism resulting in the formation of cell clusters,

while several models have been proposed to describe this process (Ratcliff et al. 2012, Gao et al. 2019, Colizzi et al. 2020, Fisher et al. 2020). A vital aspect of this would be the division of labor, through which cells specialize in performing distinct tasks while being co-dependent to other cells (Cooper and West 2018). Eventually, the high diversification towards a distinct function led to diversification at a gene expression level. The possibility of task redistribution to many different cells and cell types allowed the first organisms to survive through the changing environment, redistribute the energy costs and gave natural selection a diverse substrate for the evolutionary process to occur. The resulting multicellular organisms consist of specialized cell types, tissues, and organs that guide all aspects of their life cycle. Each cell type consists of cells with similar morphological and molecular features that can perform distinct tasks. Some of the most common tasks among all animal taxa are communication, contraction, secretion, digestion, and movement.

The presence of cell types performing such similar tasks across animals raises questions on how similar these cell types are and whether they have common evolutionary origins, questions for which we still do not have the answer. However, during the last decade, technological advances that allow a more in-depth than ever characterization of cell types might provide us with an answer.

1.4. Gene Regulatory Networks (GRNs)

Gene expression is controlled by genetic elements and gene product interactions at a specific time and place, a process known as gene regulation. This process is continuous, guiding all aspects of the embryonic development and adult life of all organisms, as well as their response to environmental changes (Davidson 2006). At the very early stages of embryonic development, gene regulation is mostly established through maternally deposited RNA molecules and proteins, ensuring that the initial developmental cleavages and specification processes occur in the most energy-efficient way. After a few cleavages, zygotic gene expression is initiated, taking over this process and resulting in distinct chromatin states and spatiotemporal expression profiles of gene products. The main difference between the two gene regulatory modes is that maternal gene regulation depends mainly on post-transcriptional mechanisms, while zygotic gene regulation uses both

transcriptional and post-transcriptional mechanisms (Shen-Orr et al. 2010). Regardless of the mode used, both are crucial for embryogenesis and result in differential gene expression across cell types. The total genes expressed in a cell or cell type and the regulatory mechanisms controlling their expression constitute its molecular signature/fingerprint.

Therefore, to understand how diverse cell types are established as well as their function, it is necessary to identify which genes are expressed within them, when they are expressed, and which are the genetic elements controlling their expression. Such interactions are depicted in Gene Regulatory Networks (GRNs), which are hierarchical logical maps of inputs and outputs (Fig. 1.1) encoding for genes active in a specific cell or cell type at a given time point (Levine and Davidson 2005). GRNs control the establishment and maintenance of a cell type's identity, and they include subsets of genes encoding for transcription factors (TFs), the main drivers of gene regulation, signaling molecules (ligands, receptors, and secondary messengers), and terminal differentiation genes that reflect the cell type's role (Levine and Davidson 2005). Several GRNs have been already described in various organisms spanning from plants to animals (Krouk et al. 2013) and have been used to understand the intersection between genome and development (Arnone and Davidson 1997).

The emergence of multicellularity and diversification of cell types is linked to the increased complexity of gene regulatory networks (Chen and Rajewsky 2007). Therefore, identifying the GRNs operating within cell types across species allows cell type comparisons at a regulatory level that can reveal the evolutionary history of a given cell type and assist in disentangling its function (Davidson and Erwin 2006). For instance, such comparisons can reveal whether similar GRNs controlling the morphogenesis of one feature are redeployed (co-opted) in new developmental contexts (McQueen and Rebeiz 2020). On the other hand, studying GRN evolution can reveal novel GRN modifications at different regulatory levels that can lead to new body plans and speciation (Davidson and Erwin 2006).

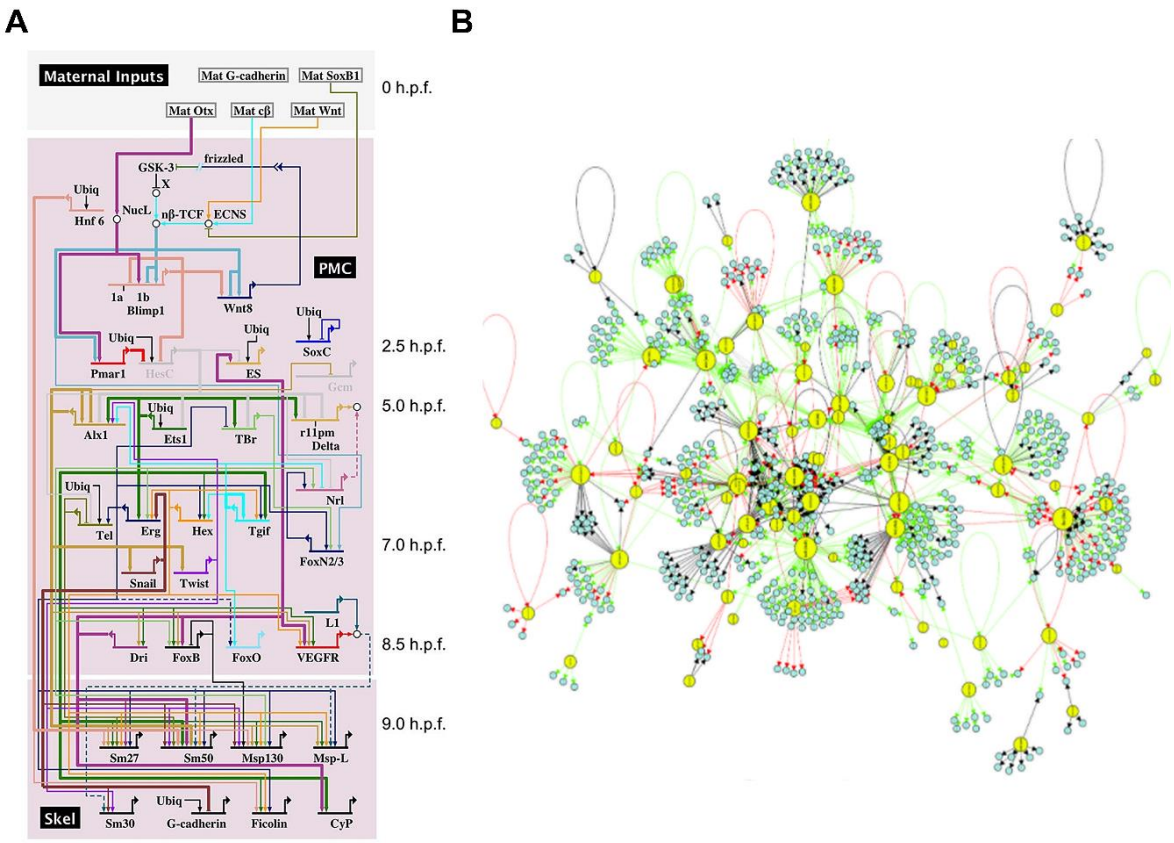


Figure 1.1. Different representations of GRNs. A) GRN visualization in which the regulatory region of each gene is represented by a horizontal bar, and the gene's name is written under the bar. If the gene's output (line with the same color) ends with an arrowhead, the transcription factor activates gene expression. If the output ends with a line, the transcription factor suppresses gene expression. B) GRN visualization in which regulatory genes are represented as yellow dots and target genes as blue circles. Black lines indicate an unknown mode of regulation; the green line indicates that a gene act as an activator; red lines indicate repression, and grey lines indicate a dual regulatory mode (Saunders and McClay 2014, Medeiros Filho et al. 2019).

1.5. Tools to unravel cell type identity

Unraveling the molecular signature and wiring of a cell type is crucial for understanding its identity, function, and evolutionary origins. There are different levels from which this information can be obtained, spanning from revealing the chromatin organization, gene product interactions, regulatory elements topology, and the amount of genes expressed in a cell or cell type at a specific time, to describing the 3D dimensional organization of protein structures, interactions, and functions.

Extensive effort has been taken on understanding the molecular signature and wiring of cell types across taxa. Most of our current knowledge on cell type identity comes from traditional molecular biology approaches such as identifying molecular

markers (mRNA molecules or proteins present), fate mapping, and gene expression perturbation—the ability to perturb gene expression allowed deciphering several mechanisms controlling differential gene expression across cell types. Gene perturbation can be lead to either silencing the expression of a gene (knocking-down or knocking-out) or overexpressing it either in its regular expression domain or ectopic ones (knock-in). Recent advances in sequencing, microfluidics, and nucleic acid barcoding have allowed the throughput recognition of cell types molecular signatures even at a single cell level (Kashima et al. 2020).

Technologies such as Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) and the recently developed single cell ATAC sequencing (scATAC-seq) have been extensively used to understand which parts of the chromatin are active (gene expression is taking place) at a bulk organismal or a single cell level respectively (Magri et al. 2019, Baek and Lee 2020).

Single cell RNA sequencing (scRNA-seq) technology developed during the last decade is an innovative method used for unraveling the transcriptional content of individual cells, while transcriptionally similar single cells are computationally grouped, forming clusters that represent cell types. ScRNA-seq is able to produce single cell transcriptomes reflecting the cell to cell transcriptomic variability, which is usually hidden by bulk transcriptomic strategies that highlight the average expression of genes in a group of cells consisting one or more cell types. Therefore, the primary advantage of this method is that it simultaneously produces thousands of single cell transcriptomes that can reveal cell type-specific molecular signatures. Several scRNA-seq systems have been developed that mainly differ in how the mRNA transcripts originating from individual cells are amplified to generate either a full-length cDNA or cDNA with a unique molecular identifier (UMI) incorporated at either the 5’ or 3‘end (Baran-Gale et al. 2018). All the single cell RNA sequencing technologies rely on the partition of a cell suspension into nanoliter-scaled reactions contained in either plates, micro-wells or in water in oil droplets. For instance, SMART-seq (Switching Mechanism at 5′ End of RNA Template) is a plate based approach that provides full length transcript coverage and is frequently used to identify transcript variant isoforms as well as alternative splicing events. On the other hand, Drop-seq, InDrop (Indexing Droplets) and 10x Genomics are droplet based protocols designed to incorporate UMIs into the cDNA (Baran-Gale et al. 2018).

The workflow of all droplet-based technologies includes the dissociation of an organism, organ, or tissue into single cells, isolation and encapsulation of the single cells into droplets, barcoding of individual mRNAs, reverse transcription followed by the amplification of the resulting cDNA and sequence of the constructed cDNA library. While Drop-seq, InDrop and 10x Genomics technologies are all based on the encapsulation of single cells into water-oil droplets, they differ in how the barcoding occurs (Solomon et al. 2019). In detail, the low efficiency Drop-seq system is based on the introduction of a solid bead into a water-in-oil droplet, while InDrop system droplets contain barcoded hydrogel spheres with cells and reverse transcription reagents together allowing the incorporation of the reverse transcription regents into the cell and bead at the time of encapsulation. On the other hand, 10x Genomics, uses Gel Beads containing barcoded oligonucleotides that are combined with the single cells and water in oil on a microfluidic chip to form reaction droplets (Gel Beads in Emulsion, or GEMs). Each GEM contains a single cell, a single Gel Bead, and the reagents used for reverse transcription. Within each GEM droplet, a single cell is lysed, the Gel Bead is dissolved to free the identically barcoded reverse transcription oligonucleotides into solution, and reverse transcription of polyadenylated mRNA takes place. Once, reverse transcription is complete all the cDNAs from a single cell have the same barcode, allowing the sequencing reads to be computationally mapped back to their single cells they originated from.

Since the first implementation of scRNA-seq in 2009 used to describe distinct cell types during early mouse embryonic development, scRNA-seq has been extensively used to characterize cell types in a variety of species and several experimental conditions. This technology has been used to identify distinct molecular signatures, revealing novel cell types in a variety of taxa ranging from insects (Davie et al. 2018, Severo et al. 2018, Cho et al. 2020), cnidarians (Sebe-Pedros et al. 2018), and sea squirts (Cao et al. 2019, Sharma et al. 2019) to more complex ones such as mice (Nestorowa et al. 2016, Jung et al. 2019, Ximerakis et al. 2019, Yu et al. 2019, Qi et al. 2020) and even humans (Yu et al. 2019, Esaulova et al. 2020, Qi et al. 2020, Zhao et al. 2020). For instance, scRNA-seq of the sea anemone *Nematostella vectensis* led to the identification of pharyngeal cells that give rise to both neurons and digestive cells (Steinmetz et al., 2017), as well as to shared molecular fingerprints between neuronal and secretory gland cells in

cnidarians (Sebe-Pedros et al. 2018). Moreover, single cell RNA-sequencing technology has been sufficient to describe time-sensitive and complex processes such as cell type specification and differentiation leading in many cases to the successful lineage tracing of complex cell types to their progenitor cells by reconstructing developmental trajectories (Wagner et al 2018, Cao et al. 2019, Cho et al. 2020).

Lastly, scRNA-seq has been recently used to perform comparative cross-species studies which led to the identification of similar cell types across taxa. For instance comparative scRNA-seq approaches have identified a shared molecular signature between several human and mouse bladder cell types, while at the same time recognizing two novel human bladder cell types (Yu et al. 2019).

1.6. Sea urchin as a model to study development and evolution

Sea urchins are marine invertebrates, and the embryonic development of most sea urchin species is indirect, producing an intermediate life stage, the larva. The free-swimming and feeding sea urchin larva spends several months in the water column and will metamorphose to the sea urchin juvenile in the final stage of development. The transition of the larval to the adult life cycle is dramatic and accompanied by a change of body plan symmetry from bilateral to radial (Ruppert et al. 2004). Sea urchins together with sea stars, sea cucumbers, sea lilies, and brittle stars form the phylum Echinodermata (Fig. 1.2).

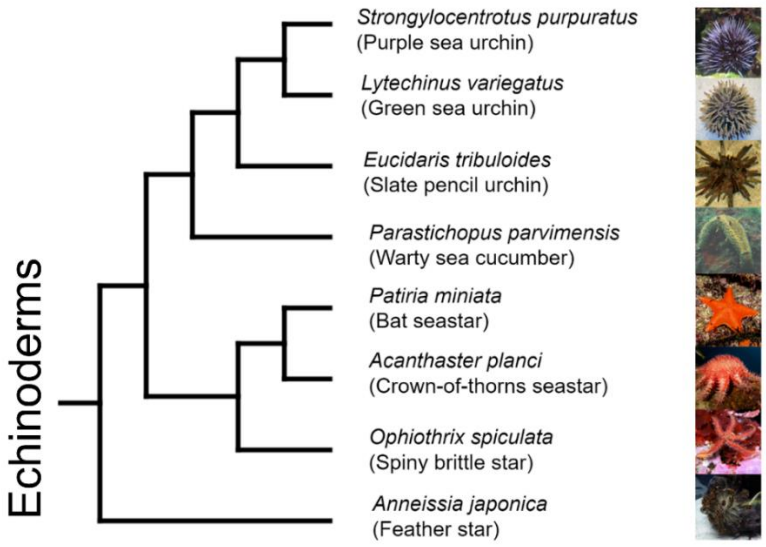


Figure 1.2. Representative members of the echinoderm clade. Figure adapted from Echinobase (https://new.echinobase.org).

Due to their phylogenetic position as an early branching non-chordate deuterostome, echinoderms are an ideal model to study the evolution of different cell types among deuterostomes. Apart from evolutionary studies, sea urchins and especially sea urchin embryos, have been extensively used to describe fundamental cellular and developmental biology concepts such as cell division, cell migration, cell specification/differentiation, and gastrulation throughout embryonic and larval development.

In detail, sea urchin has been one of the first model systems used in developmental biology starting from 1847 when Derbes, Dufosse, and von Baer took advantage of the sea urchin embryos' transparency to observe the fertilization process (Briggs and Wessel 2006). From that point and onwards, many scientific breakthroughs followed, with some of the most important as summarized by Ernst (2011) being:

- The fertilization of an oocyte by only one sperm cell.
- The observation that chromosomes contain the genetic information about different traits.
- The discovery of DNA localization in chromosomes residing in the nucleus and that RNA resides in the cytoplasm.
- The first correlation between RNA and protein synthesis.
- The finding that maternal mRNAs are placed into the egg during oogenesis.

Almost 200 years after those breakthroughs, sea urchins are still a powerful model system to study fundamental developmental processes such as cell type specification and differentiation because of the many advantages they have, some of which are mentioned below.

Different sea urchin species are widely spread across the seas worldwide, and the collection of adult individuals is relatively easy. Adult male and female individuals can produce millions of gametes (oocytes and sperm) that they release in the water column, where fertilization occurs (external mode of fertilization). Such a high number of gametes leads to an equally high number of fast and synchronous developing embryos and larvae, thus a high amount of experimental material.

An important factor promoting the establishment of sea urchin embryos as a developmental biology model system is the ease with which molecules can be introduced into sea urchin embryos via microinjection (Cheers and Ettensohn 2004). Examples of this are the microinjection of morpholino antisense oligonucleotides to

silence genes, synthetic mRNAs and proteins resulting in overexpression, and cis-regulatory modules to decipher regulatory mechanisms (Flytzanis et al. 1985, Arnone et al. 2004, Materna 2017). As of recently, gene editing is also possible by implementing the Crispr/Cas9 editing system that allows the knock-in/out of genes (Lin and Su 2016, Oulhen and Wessel 2016, Lin et al. 2019, Liu et al. 2019, Wessel et al. 2020, Yaguchi et al. 2020).

A reason for the extensive use of sea urchins as a model for evolutionary studies is the simplicity of their genome compared to vertebrates. For instance, the genome of the sea urchin *Strongylocentrotus purpuratus* contains around 23,000 genes (Sea Urchin Genome Sequencing et al. 2006), a number comparable to the approximately 15,682 and 24,000 genes encoded in the *Drosophila melanogaster* and human genomes respectively (International Human Genome Sequencing et al. 2004, Hoskins et al. 2015). The main difference lies in the fact that the echinoderm and vertebrate lineages branched out before large scale genome duplication events. This means that many of the genes found as multiple paralogues in vertebrates correspond to a single sea urchin homolog, enabling the study of complex developmental processes and regulatory interactions on an organism that bares single gene copies. Moreover, several sea urchin species genomes have already been sequenced offering the possibility of performing microevolutionary genetic comparisons, thus leading to the discovery of novel gene regulatory wirings and further understanding the origins of novel cell types (Kinjo et al. 2018, Davidson et al. 2020).

1.7. Sea urchin embryonic and larval development

Sea urchins are gonochoristic animals that reproduce sexually via releasing female and male gametes into the water column, leaving reproduction to chance. Once fertilization occurs, a fertilization envelop is formed, inhibiting further fertilization by more than one sperm cell. From this point and onwards, the embryonic development is initiated, and the zygote goes through a series of rapid cleavages giving rise to a ciliated blastula. At this developmental stage, the embryo consists of a monolayered epithelium forming an empty sphere (blastocoel).

The formation of the first opening of the future digestive tract at the center of the vegetal plate, the blastopore, marks the beginning of gastrulation. Once the blastopore is formed, the forming archenteron elongates within the blastocoel towards the anterior domain. As soon as the archenteron is fully elongated, several morphogenetic processes take place. Cells located at the tip of the archenteron migrate to both sides of the anterior domain of the archenteron forming two bilaterally symmetric coelomic pouches, while another population of cells located at the same area interacts with the neighboring oral ectoderm cells (stomodeum) to form the second opening of the digestive tract; the mouth. Once the mouth opening is formed, embryogenesis is concluded, and larval development begins. At this stage, the larva is equipped with a ventrally positioned ciliary band, and its body shape changes into a cone-shaped echinopluteus, while additional ciliated structures, the larval arms, are formed in the oral and aboral edges of the ciliary band. The arms are supported by endogenous calcite rods and are involved in capturing phytoplankton via creating a larger ciliated surface that can create additional water flow directed towards the larva's mouth. The free-swimming and feeding larva's digestive tract is compartmentalized into distinct domains due to the formation of sphincter structures (esophagus, stomach, intestine) and equipped with a musculature that regulated the feeding process. As development goes on, the ciliary band is restricted to structures resembling literal bands of cells, more arms are formed, and the sea urchin rudiment starts to develop on the left side of the larva. Upon induction by environmental cues, competent larvae with a fully developed rudiment settle and undergo metamorphosis. Sea urchin development is a relatively fast process, the timing of which varies across sea urchin species, but in all cases is concluded within several months. In the sea urchin *S. purpuratus,* embryogenesis is a three-day process followed by three months of larval development, and an overview of the life-cycle is presented in Figure 1.3.

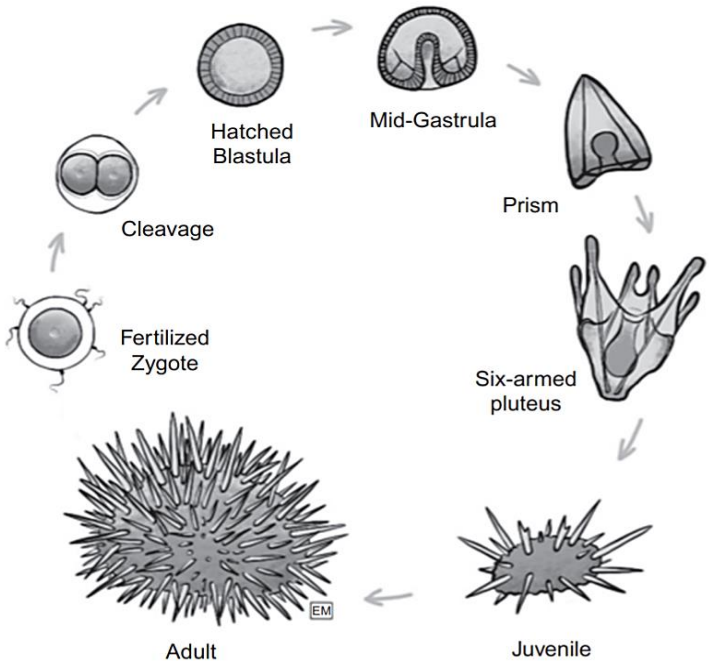


Figure 1.3. Sea urchin's life-cycle. Schematic representation of sea urchin embryonic and larval development starting from the fertilized egg until the adult sea urchin (Figure from Adams et al. 2019).

1.8. Sea urchin cell type specification

Sea urchin is a triploblastic organism, and the specific cells that constitute the three germ layers (mesoderm, ectoderm, and endoderm) are specified by the 4th cleavage (16 cell stage), with the animal pole cells giving rise to ectoderm and the vegetal pole cells to endomesoderm (Gilbert and Barresi 2018). The cell type specification process has been studied in great detail, starting from the zygote until the early pluteus stage, a developmental stage in which embryonic development is completed, and the larval lifestyle begins. The initial steps of embryonic development rely on maternal inputs (RNAs and proteins) deposited in the egg that guide early development until the zygotic expression is initiated, approximately after the 10th cleavage (mid-blastula stage) (Kipryushina and Yakovlev 2020). In most cases described so far, silencing the expression of such maternal factors has a severe effect on embryogenesis, highlighting how fragile the early embryonic developmental program is (Range and Lepage 2011, Yaguchi et al. 2018, Tsironis et al. 2021). These maternal inputs are necessary to kick-start the various gene regulatory networks and subsequent cell specification and differentiation cascades.

Events resulting in body plan asymmetry and establishing the embryonic and larval axes coincide as cell type specification. One of the earliest events during early embryogenesis is the formation of the small micromeres during the 4th cleavage, a set-aside cell population, smaller in size than the rest of the embryonic cells, resulting in an asymmetry at the vegetal pole. This population that remains dormant until the late gastrula stage and later starts to proliferate and migrate to the coelomic pouches of the larva has been long hypothesized to be involved in the germline and juvenile sea urchin formation (Pehrson and Cohen 1986, Juliano et al. 2010, Yajima and Wessel 2011). At the same developmental stage, the primary anterior/posterior axis is being established due to Wnt signaling being activated at the vegetal pole of the embryo. Wnt signaling activity results in an uneven distribution of b-catenin in this domain's most posterior cells and gives rise to the endomesoderm germ layers (Wikramanayake et al. 1998, Logan et al. 1999). An immediate result of this signaling activity is the establishment of the anterior/posterior axis, in which the most posterior part is fated to give rise to endomesoderm and the most anterior part to ectodermally derived cell types (Fig. 1.5). Moreover, Wnt signaling is involved in the spatial restriction of one of the embryo's neurogenic domains; the anterior neuroectoderm (ANE). The mechanism through which Wnt signaling exhibits this role includes the repression of neurogenic transcription factors in all other embryonic domains except the anterior neuroectoderm, a region in which opposing signals prevent Wnt signaling from being active (Range et al. 2013, Yaguchi et al. 2016, Range and Wei 2017, Martinez-Bartolome and Range 2019).

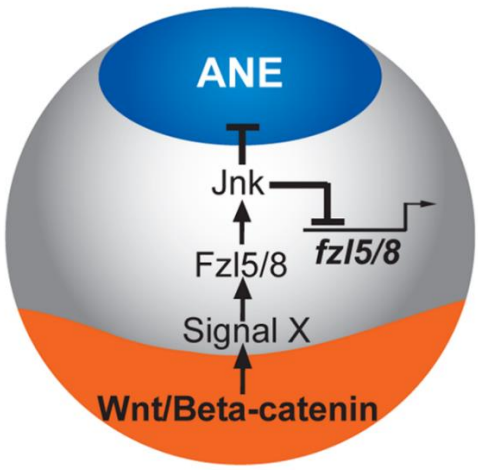


Figure 1.4. Establishment of anterior/posterior axis. Wnt signaling mediated specification of vegetal pole and restriction of ANE. (Figure adapted from Range et al. 2013).

The next embryonic axis to be specified is the ventral/dorsal or oral/aboral axis. Maternal inputs and uneven distribution of mitochondria in the zygote initiate this process, leading to differential expression of several Tgf-β family members, whose gene products and induced signalings are active in different embryonic domains. During cleavages, one embryonic domain inherits a higher amount of mitochondria and is specified as ventral ectoderm (Coffman et al. 2004). The first morphological evidence of bilateral symmetry resulting from the establishment of the ventral/dorsal axis is the even distribution of two clusters of Primary Mesenchyme Cells (PMCs), the skeletogenic cells of the embryo, next to the archenteron at the gastrula stage. On the other hand, the first molecular evidence has been detected much earlier at the 60-cells stage embryo. At this stage, the Tgf-β family ligand *Nodal* starts to be expressed due to the mitochondria asymmetry, activating the ventral ectoderm gene regulatory network (Duboc et al. 2004).

On the other hand, *Nodal* also activates the expression of the Tgf-β family members *Lefty*, *Bmp2/4,* and *Chordin*, which are responsible for the establishment of distinct domains along the ventral/dorsal axis (Duboc et al. 2004, Lapraz et al. 2009, Duboc et al. 2010, Saudemont et al. 2010, Lapraz et al. 2015). For instance, Bmp2/4 protein is active in the ectodermal domain fated to give rise to dorsal ectoderm. In contrast, Lefty and Chordin proteins are responsible for inhibiting Nodal and Bmp2/4 signalings respectively from being active in the ectodermal domain that is fated to give rise to the ciliary band, through antagonizing Nodal's and Bmp2/4's binding to their receptors, respectively (Duboc et al. 2004, Duboc et al. 2008, Saudemont et al. 2010). The establishment of both anterior/posterior and ventral/dorsal axes also result in the restriction of the neurogenic capacity of the embryo to three neurogenic regions of the embryo, the anterior neuroectoderm (ANE), the ciliary band (CB), and the neurogenic part of the endoderm (Angerer et al. 2011).

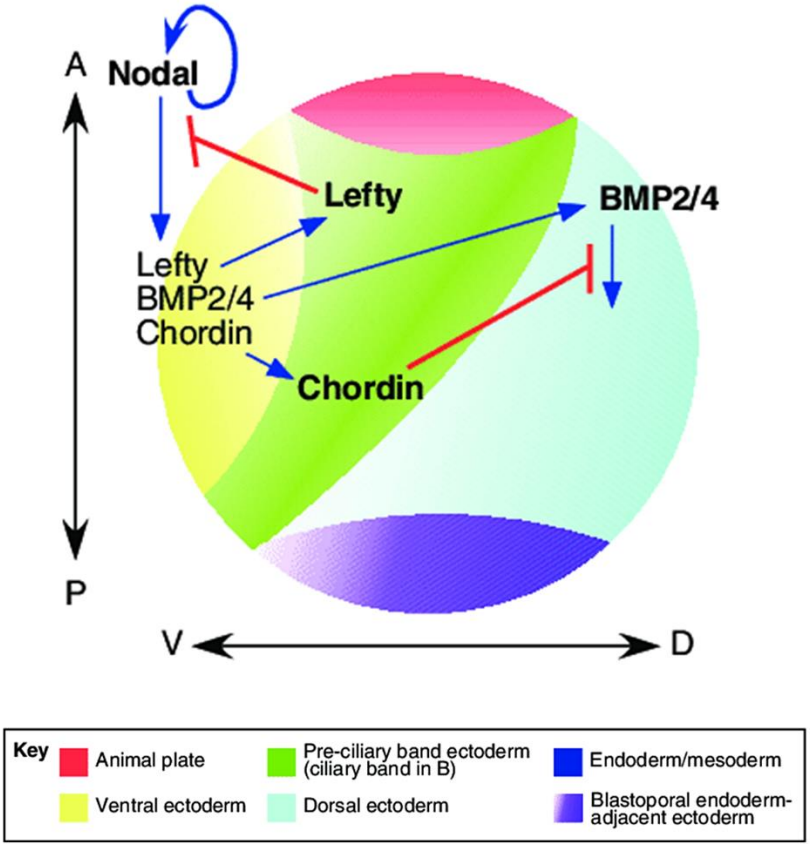


Figure 1.5. Establishment of ventral/dorsal axis. Schematic representation of the Tgf-β family signalings involved in establishing the ventral/dorsal axis. A, anterior; D, dorsal; P, posterior; V, ventral. (Figure adapted from Angerer et al. 2011)

1.9. Sea urchin cell type differentiation

By the gastrula stage, most sea urchin cell types have already been specified, and differentiation occurs. The differentiation process is continuous and leads to the formation of a planktonic pluteus larva equipped with the necessary machinery to feed and grow until the sea urchin rudiment is well-formed and metamorphosis begins. Thanks to extensive traditional experimental work and lineage tracing experiments, most of the embryonic cell types and their origins have been identified and mapped onto distinct domains (Fig. 1.6). Within the developmental window of 2 to 3-days post-fertilization, differentiation of significant cell types occurs, most of which are involved in the feeding process and immune defense.

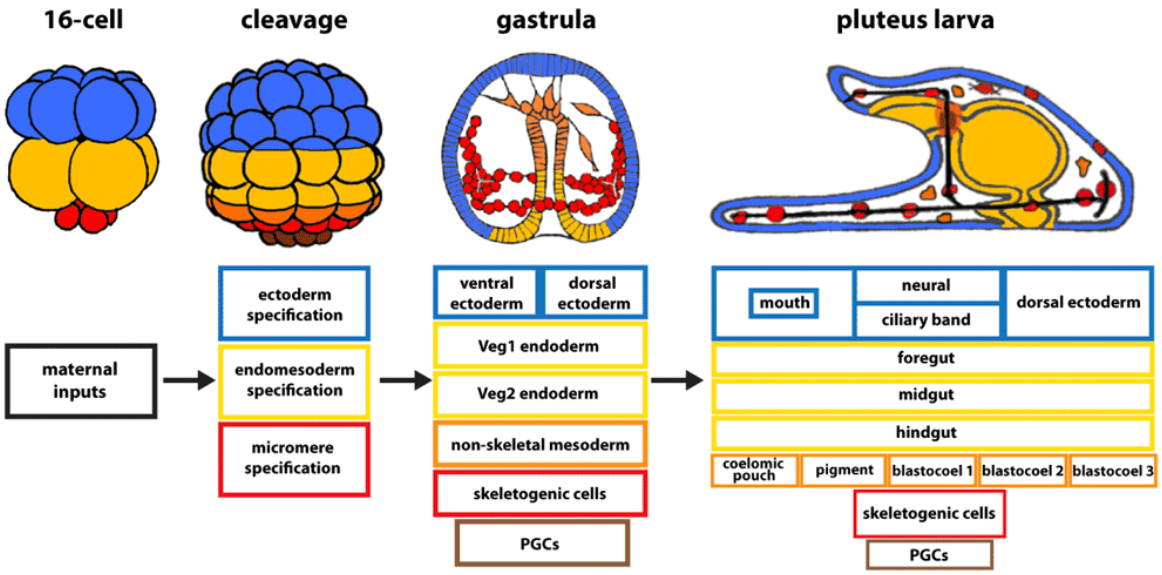


Figure 1.6. Cell type specification and differentiation. Schematic representation depicting the developmental origins of the larval cell types. PGCs, Primordial germ cells. (Figure adapted from Martik et al. 2016).

1.9.1. Derivatives from the endoderm

Morphological and molecular studies have identified distinct endodermal precursors that differentiate to various cell types by the pluteus stage, and the GRN controlling this process has been reconstructed to a great level (Annunziata and Arnone 2014). The archenteron of a gastrula morphologically resembles a straight tube open on one side (blastopore) that develops into a tripartite gut by the pluteus stage. The most anterior part is the foregut, the most posterior the hindgut, and the one in between the midgut. These three different domains give rise to specific cell types that form the pluteus' digestive tract. Foregut consists of a diverse pool of precursors. Molecular studies have identified specific foregut progenitors expressing the neurogenic transcription factors *Brn1/2/4, Six3,* and *Nkx3.2* (Cole and Arnone 2009, Wei et al. 2011). These neuronal precursors give rise to two differentiated neuronal populations placed near the larva's mouth (Wei et al. 2011, Wood et al. 2018). Most of the remaining foregut precursors contribute to the esophagus formation, whereas those at the very tip interact with ectodermal cells to form the larva's mouth. On the other hand, endodermal cells placed at the foregut and midgut border differentiate to form one of the two sphincter cells; the cardiac sphincter (Andrikou et al. 2013). These cells control the water flow that contains food from the mouth and the esophagus to the rest of the digestive system.

The midgut domain is destined to differentiate into the larva's stomach and complex cell types involved in food digestion, and the molecular fingerprint of most midgut precursors includes the expression of *Endo16, ManrC1A,* and *Chip* (Annunziata et al. 2014). A small population of two bilaterally symmetric cells located in the anterior part of the midgut gives rise to a unique digestive domain referred to as exocrine pancreas-like cells (Annunziata et al. 2014, Perillo et al. 2016). These cells express genes encoding for the transcription factors *Hnf1a* and *Ptf1a* and enzymes involved in food digestion such as carboxypeptidases, lipases, and amylases and have a genetic program similar to the vertebrate pancreatic acinar cells (Perillo et al. 2016). Cells at the border of midgut and hindgut domains that express the transcription factors *Pdx1* and *Nkx6.1* differentiate at the pluteus stage into another type of sphincter cells; the pyloric sphincter (Cole et al. 2009, Israel et al. 2016). The pyloric sphincter controls the digested food flow from the stomach to the digestive system's final components, where the non-digested parts are discarded.

The hindgut is a diverse endodermal domain that contains precursors that give rise to various cell subtypes. Much evidence suggests that the hindgut already at the gastrula stage consists of at least two different subdomains. The largest part of the hindgut consists of *Pdx1*, *Cdx*, *FoxI,* and *FoxD* positive cells, whereas cells residing close to the blastopore express the specific markers *Wnt10*, *Cdx*, *Bra,* and *Hox11/13b* (Annunziata et al. 2014). The cells that compose the first hindgut domain form the intestine of the larva, whereas the second one differentiates to form the last part of the digestive tract, the anus. Recent studies showed that at a late pluteus stage, the posterior gut domain deriving from the hindgut also supports neurogenesis and the formation of nitric oxide producing neurons (Yaguchi and Yaguchi 2019).

1.9.2. Derivatives from the mesoderm

Mesodermal diversification is one of the earliest events taking place during the 5th embryonic cleavage. At this stage, two mesodermal domains branch out; the skeletogenic mesoderm (SM) and the non-skeletogenic mesoderm (NSM). Skeletogenic mesoderm is a relatively uniform cell type consisting of PMCs, whereas NSM consists of a pool of various precursors destined to form different types (Sethi et al. 2009, Sharma and Ettensohn 2011). Notably, the mechanisms and gene regulatory cascades taking place during SM specification and

differentiation into specific skeletal populations have been identified at a gene regulatory level (Shashikant et al. 2018).

At gastrula stage, PMCs are distributed in a symmetrical pattern bilaterally to the archenteron and arranged in a semicircular way across the dorsal ectoderm. PMCs extend filopodial protrusions, congregate and fuse to form a syncytium (Hodor and Ettensohn 1998, Sethi et al. 2009, Sharma and Ettensohn 2011). The spatial organization of the syncytium is controlled by ectodermal cues, including the Vegf and Fgf signalings that ensure the correct spatial location of the magnesium calcite skeletal rods (Guss and Ettensohn 1997, McIntyre et al. 2014).

Non-skeletogenic mesodermal precursors that give rise to diverse cell types are already in place at the gastrula stage, and they are among the first cell types to enter the blastocoel during gastrulation (Andrikou et al. 2013). Once gastrulation is initiated, NSM cells are entering the blastocoel attached to endodermal cells. As gastrulation continues, NSM progenitors remain attached to the forming foregut's tip until this process is concluded. Small micromeres are members of this group of diverse cells as revealed and express the specific markers *Vasa, Nanos,* and *Seawi* (Juliano et al. 2006, Yajima et al. 2014). Small micromeres stay dormant, and in the developmental window from gastrula to pluteus stage, start to proliferate. Then they migrate, and along with ectodermal tissues, they constitute the coelomic pouches of the larva. The left coelomic pouch develops into the rudiment, harboring the growing juvenile sea urchin.

Another NSM derivative is the group of blastocoelar cells and is among the first NSM cells to migrate from the tip of the foregut to pattern the blastocoelar space. Blastocoelar cells form networks of interconnected filopodial cells and are believed to be involved in the larva's immune response (Ho et al. 2017). The NSM pool of cells contains two additional immune system cell types; the globular and the pigment cells. Globular cells are spherical shaped cells first identified by the expression of the perforin-like enzyme MacpfA2 (Ho et al. 2017). They are scattered throughout the blastocoel, and they have a surveillance role (Ho et al. 2017). Pigment cells are red/orange cells scattered throughout the larval ectoderm, especially in the aboral ectoderm layer. The pigment they contain is a naphthoquinone called echinochrome and is accumulated in pigment cell precursors during gastrulation. Upon infection, pigment cells migrate from the ectoderm to the blastocoel and initiate an immune response upon communication with other cell types (Ho et al. 2017). Gene

perturbations that affect the pigmentation pathway lead to albinism and animals with reduced response to environmental changes (Wessel et al. 2020).

Specified muscle progenitors are found at gastrula stage, and at the pluteus stage, they differentiate to form circumesophageal apparatus (Andrikou et al. 2013). Esophageal muscle fibers are patterning the esophagus and the digestive tract's sphincters, ensuring the rhythmical contraction and flow of water containing phytoplankton (Annunziata et al. 2014, Yaguchi et al. 2017). Disruption of the musculature formation leads to larvae that are unable to feed, and eventually, they perish.

1.9.3. Derivatives from the ectoderm

Ectodermally derived progenitors differentiate towards mostly epithelial tissues and neurons constituting the larva's nervous system. At gastrula stage, the oral ectoderm domain, also known as stomodeum, is an epithelial tissue placed in the ventral region where the mouth will form, and the molecular signature of the stomodeum includes the transcription factors *FoxA, Gsc,* and *Bra* (Saudemont et al. 2010). During mouth formation and opening at approximately 3 dpf, stomodeum and foregut cells collaborate and intermingle to ensure the invagination of the foregut into the stomodeum domain and opening of the mouth. On the dorsal side, close to the blastopore, aboral ectoderm epithelial cells reside. At larval stages, this domain elongates to form a cone-shaped structure. The role of these cells remains unknown, although they seem to participate in immune response as they are harboring several immune cell types.

Almost the entire body of the developing embryo and larva is covered by ciliated epithelial cells that allow the animal to move. During the pluteus stages, many ciliated cells arrange in bands forming the ciliary band of the larva, while several genes have been found to regulate this process, including the transcription factors *Hnf6, Dri,* and *FoxG.* (Barsi et al. 2015).

One major ectodermally derived cell type is the anterior neuroectoderm placed in the most apical and dorsal part of the embryo. At gastrula stage, ANE is a thickened epithelium containing ciliated cells with longer cilia than the ciliary band constituting the apical tuft. At pluteus stage, the morphology of ANE changes into a plate-like structure (apical plate) that contains the larva's chemosensory center (apical organ) and several neuronal cell types (Marlow et al. 2014).

The ciliary band and apical plate together are the two major neurogenic regions of the larva. Apical plate and predominantly apical organ neurons, as well as the ciliary band neurons, are considered the central and peripheral nervous systems of the larva, respectively, and their interconnection guides every aspect of the larva's life (Burke 1978, Bisgrove and Burke 1987, Burke et al. 2006, Hinman and Burke 2018).

1.10. Sea urchin nervous system organization

The first neurons of the sea urchin embryo arise at the gastrula stage. The pre-larval nervous system consists of mostly non-differentiated neuronal precursors scattered across the forming ciliary band and ANE as well as of endodermally derived foregut neuronal precursors. ANE gives rise to the apical organ (AO), a thickened neuroepithelium placed in the most anterior part of the larva that is considered to be the central nervous system of the larva. This centralization theory is based on several observations such as the expression of a subset of genes whose vertebrate homologs are wiring the forebrain region, the utilization of similar specification mechanisms (Wei et al. 2009, Range 2014), and the directionality of the axonal projections from the ciliary band neurons towards the apical organ (Burke et al. 2014). At gastrula stage, only two differentiated neuronal types have been described so far; the sensory serotonergic neurons in the ANE region and two cells in the oral ectoderm that differentiate into neurons patterning the post-oral arms of the larva (Angerer et al. 2011, Garner et al. 2016).

Nonetheless, the differentiation cascade of these neuronal types has been studied in great detail, and it has been demonstrated that all neuronal precursors identified so far go through a similar stepwise differentiation process. Interestingly, lateral inhibition mediated by Delta/Notch signaling, known to be a defining step in the neurogenic mechanism across taxa (Lasky and Wu 2005), has also been found to be active during sea urchin neurogenesis (Mellott et al. 2017). For all described neuronal types so far, the next step of the cascade includes a transient expression of the transcription factor *SoxC* followed by the expression of the transcription factor *Brn1/2/4*. After this crucial step, their differentiation mechanisms diversify, and the neuronal precursors express distinct regulator programs (Garner et al. 2016).

The nervous system of the early pluteus larva consists of 40-50 neurons patterning distinct domains of the larva that correspond to the three neurogenic domains

mentioned above (Wei et al. 2016). Traditionally, the classification of the neurons is based on the analysis of the neurotransmitters they produce, such as catecholamines (dopamine, adrenalin, and noradrenalin), indolamines (serotonin), and amino acids (histamine, acetylcholine, and GABA) (reviewed in D'Aniello et al. 2020). Such analyses identified serotonergic neurons in the apical organ, dopaminergic post-oral neurons, and cholinergic ciliary band neurons (Katow et al. 2010, Adams et al. 2011, Wei et al. 2016, Slota and McClay 2018).

A recent study revealed that the nervous system of the early pluteus larva is far more complex than previously believed. The authors of this study re-established the traditional classification by taking into account another class of neuromodulators: the neuropeptides. Neuropeptides are secreted signaling molecules engaged in many physiological functions of the nervous system in most metazoans that function through specific G-protein coupled receptors (GPCRs) to control neuronal regulation, physiological and behavioral processes (Elphick et al. 2018, Jekely et al. 2018). Thorough characterization of the expression patterns of such neuropeptides, combined with co-expression analyses of known neurotransmitters and neuronal markers, revealed the presence of seven differentiated neuronal populations the early pluteus larva. In detail, one neuronal population was found to be present in the apical plate, one distally and close to it, two adjacent to the larva's mouth, one in the stomach region, one in the post-oral arms, and one laterally to the ciliary band (Wood et al. 2018).

Although much effort has been taken to unravel the neuronal types that pattern the pre-larval and early larval nervous system, little is known about their function. Numerous studies have identified the roles of several biogenic amines in the control of movement and growth, with serotonin, dopamine, acetylcholine, adrenalin, and noradrenaline to be involved in regulating the ciliary beating (reviewed in Marinkovic et al. 2020). Furthermore, Adams and colleagues found an additional role of dopaminergic signaling in regulating a phenotypic response of larvae linked to food availability. When food is absent, fastened larvae grow longer arms than fed larvae and thus larger ciliated surface compared to the one developed by fed larvae in an attempt to produce stronger currents and drive water containing food towards their mouth. Adams et al. showed that this phenotypic response is related to algae-induced dopamine signaling inhibiting arm elongation in fed larvae. Finally, the only reports on histamine's function come from studies carried out on pre-metamorphic

larva in which the authors demonstrated that pharmacological inhibition of histamine receptors induced spontaneous settlement and metamorphosis (Sutherby et al. 2013).

Aims and goals

This thesis aims to identify the molecular signature of the major cell types in place during *S. purpuratus* late embryonic and early larval development and to provide a thorough characterization of the nervous system. To this end, a scRNA-seq pipeline was developed and used to unravel the molecular fingerprint of 2-days post fertilization (dpf) late gastrula and two larval stages (3 and 5 dpf). Combining single cell transcriptomics with traditional molecular biology techniques, various imaging methods, and gene expression perturbation experiments allowed an extensive characterization of the cell types and the identification of novel neuronal types.

The main results included in the present PhD thesis are organized in separate chapters as listed below:

Chapter 3 contains the single cell RNA sequencing cell types atlases concerning the aforementioned developmental stages, validations of the scRNA-seq predictions, and a thorough analysis of several cell types' molecular signature and regulatory wiring.

In Chapter 4, the nervous system diversity and complexity during sea urchin development are thoroughly assessed, and novel neuronal and pre-neuronal subtypes are described.

Chapter 5 is about novel neuronal and non-neuronal cell types that share a molecular signature similar to the endocrine and exocrine pancreatic cells of vertebrates. Their molecular signature, wiring, and function are presented, and a case concerning their evolution is being made.

Chapter 2

Materials and Methods

This chapter contains details on the different materials and methods used for this PhD project, spanning from adult/embryo handling, microinjection, and gene expression visualization techniques to scRNA-seq library construction and analysis.

2.1. Animal husbandry

Adults *Strongylocentrotus purpuratus* individuals were collected from the Gulf of Santa Catalina, CA, USA by Peter Halmay and distributed by Patrick Leahy (Kerckhoff Marine Laboratory, California Institute of Technology, Pasadena, CA, USA). Once obtained, the sea urchin individuals were kept in circulating seawater aquaria at Stazione Zoologica Anton Dohrn in Naples under the care of Davide Caramiello. The seawater temperature inside the aquaria was approximately 14-15 ºC, which is sufficient to prevent the adult individuals from spawning spontaneously. Adult sea urchins were fed with algae, corn, and egg white.

2.2. *In vitro* fertilization of gametes and culturing of embryos and larvae

Sea urchin female and male individuals do not bare any morphological differences that could indicate their sex. The sex can only be determined after spawning induction, and gamete release has already begun from the gonopores, with the oocytes being yellow/orange and the sperm being white (Fig. 2.1). Spawning was induced only by vigorously shaking the animals. After specimen collection, the sea urchin individuals were placed back into the seawater containing tank. The eggs were collected by placing the female (gonopores and anus) facing down on a glass beaker filled with filtered seawater (FSW). The eggs were released in the water column and were allowed to settle down by gravity. Dry sperm (concentrated) was slowly collected with a glass Pasteur pipette and placed in Eppendorf tubes. The collection has to be slow due to the high viscosity of the dry sperm in order to avoid the pipette getting clogged and losing material. Both the tube containing the sperm and the beaker containing the eggs were placed on ice to preserve them until use.
The eggs were washed with FSW once and passed through a 100 µm Nitex filter to remove coelomic fluid and sea urchin spine pieces. Eggs were fertilized in 20 ml of FSW by adding the appropriate amount of activated sperm (sperm diluted in FSW). Approximately 2 µl of dry sperm are diluted in 1/1,000 FSW and used to fertilize the eggs. This dilution factor ensures successful fertilization with lower risks of polyspermy. Since *S. purpuratus* is a Pacific Ocean native species and that the Mediterranean Sea salinity is higher than the Pacific Ocean one, FSW used for fertilization and culturing of the embryos and larvae was diluted 9:1 in deionized H_2O. In case zygotes were needed for microinjection experiments, fertilization was

performed in 2 M Para Amino Benzoic Acid (PABA) diluted in FSW to prevent the membrane's hardening. The fertilization progress was monitored under a stereomicroscope and deemed successful once the vitelline membrane was elevated. After fertilization, the zygotes were washed twice with FSW to remove the excess sperm. The zygotes were placed at 15 ^{o}C and let develop until the developmental time-points of interest. All the glassware materials used for embryonic and larval cultures were sterile and preserved from any phosphate-containing detergent that can affect normal embryonic development.

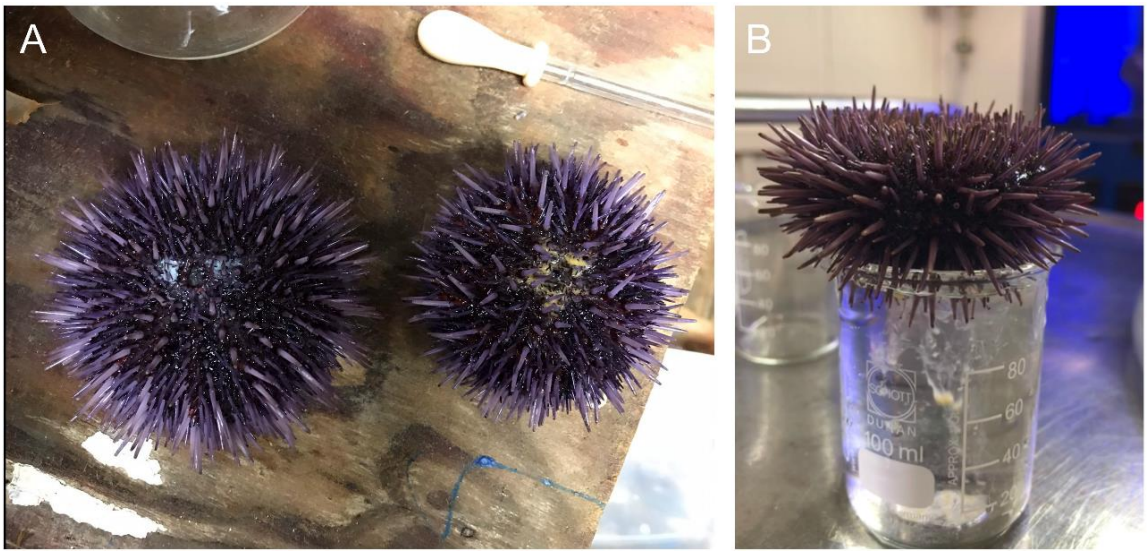


Figure 2.1. *S. purpuratus* gamete release. A) Male (left) and female (right) *S. purpuratus* individuals. Note that sexual dimorphism is absent and that sex can be determined only by the color of the gametes released. B) Collection of eggs in a glass beaker containing FSW. See how the eggs are concentrated at the bottom of the beaker due to gravity forces.

2.3. Microinjection of Morpholino oligonucleotides into sea urchin eggs

Microinjection of different constructs is an efficient way of manipulating gene expression, and microinjection protocols have been established in various organisms, with sea urchin being amongst the most effective. For this study, morpholino antisense oligonucleotides (MASOs, hereby also called morpholinos) were injected into fertilized *S. purpuratus* eggs to silence gene expression. Each morpholino used has explicitly been designed targeting the genes of interest, and their function results in blockage of the translation process. Morpholinos are *in vitro* synthesized molecules around 25-base long that differ from natural nucleic acids, and once introduced to a live cell, they hybridize to their RNA target. Unlike nucleic

acids made of ribose or deoxyribose, morpholinos are made of methylemorpholine (morpholine rings/ that replace those molecules. The presence of morpholine rings, alongside the replacement of anionic phosphates of DNA and RNA by phosphorodiamidate linkages, leads to a molecule with an uncharged backbone that is not recognized by intercellular degradation machinery (enzymes) and thus stable to nucleases. Once, injected a translation blocking morpholino hybridizes to its mRNA target. The morpholino-mRNA hybrid does not allow the molecular machinery to bind and operate, and thus, protein translation initiation is inhibited (Fig. 2.2). The remaining mRNA is slowly degrading by the mechanism of the cell.

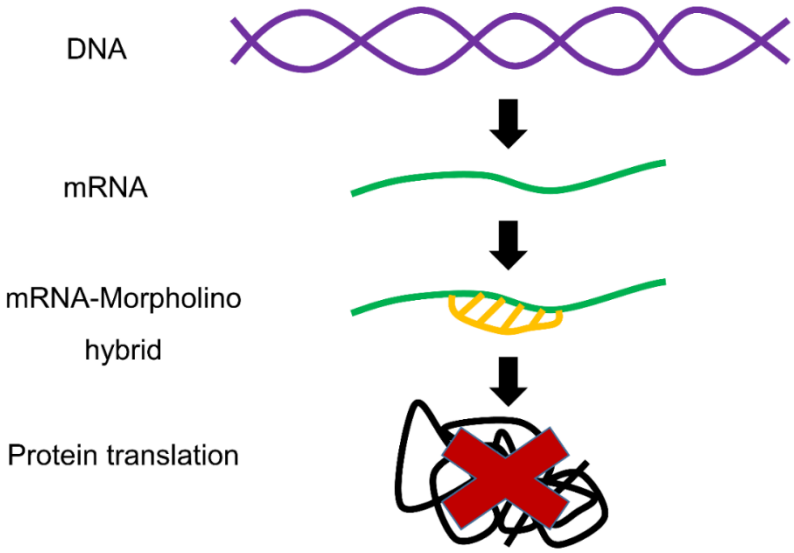


Figure 2.2. Morpholino's mechanism. Schematic representation of mechanism by which a translation blocking morpholino functions**.** The specific MASO recognizes transcribed mRNAs corresponding to a gene of interest. Morpholino is hybridized to the mRNA, blocking the initiation of translation and knocking down the gene of interest.

Microinjection procedure:

1. Morpholinos were resuspended in double-distilled H_2O (ddH_2O) according to the manufacturers' guidelines. Morpholino working solutions of 0.2 mM in 0.2 M KCl and ddH_2O were injected in fertilized eggs. All morpholinos have been acquired from Gene Tools (Corvallis, OR).
2. Morpholino working solutions were preheated for 5 min at 65 ^oC, filtered using 0.2 µm filters, and centrifuged for 15 minutes at 5,000g in a microcentrifuge. Once the solution passed through the filter, the filter was removed, and morpholino working solutions were centrifuged at maximum speed until use.

3. Morpholino working solution was transferred into a microinjection needle. Microinjection needles were prepared from 1.0 mm outside diameter, 0.75 mm inside diameter, borosilicate glass supplied by Sutter Instrument Co. Novato, CA (No. B100-75-10). Fine-tipped microinjection needles were pulled on a Sutter P-97 micropipette puller.
4. Gamete shedding and gamete collection were performed as described above.
5. A small portion of eggs was fertilized, and the fertilization process was monitored under the stereomicroscope. Zygotes were used for microinjection only when more than 95% of the eggs were successfully fertilized.
6. Fertilized eggs were transferred to a 60mm Petri dish coated with 1% agar using a fine-tipped glass Pasteur pipette, drawn out of a Bunsen flame, and broken off at the end. The pipette diameter must be the same as the diameter of the eggs, in this case, 80 µm for *S. purpuratus*.
7. Aspiration of the eggs in and out of the glass Pasteur pipette leads to dejellyed eggs and thus more permeable eggs.
8. Dejellyed eggs were placed in a row on a plastic Petri dish coated with 1% protamine sulfate. The injection dish was placed on the microinjection set up, and 2-3 drops of freshly diluted sperm in PABA containing FSW were added.
9. The tip of the morpholino containing needle was broken by adjusting the micromanipulator of the microinjection set up so that the needle gently touches the dish's surface.
10. Once the tip of the needle is broken, the morpholino working solution starts to flow. The flow and thus the amount of solution injected is controlled by the microinjection set up. The needle is quickly moved in and out of the zygote, releasing 2-4 pl of morpholino solution. The process is repeated until a few hundreds of zygotes are injected. The microinjection procedure must start immediately after fertilization since, after a few minutes, the fertilization membrane hardens and becomes impermeable.
11. Injected and uninjected zygotes are left to grow at 15 ^{o}C.

Details on material preparation for microinjection are mentioned below:

Coating Petri dishes lids with protamine sulfate: 60 mm Petri dish lids were filled with 1% protamine sulfate solution in ddH_2O and incubated for 1 minute (min). Then

protamine solution was removed, and dishes were thoroughly washed with dH_2O and left to dry until use. The protamine solution was kept at 4 ^oC and can be reused several times.

Coating Petri dishes lids with agar: 60 mm Petri dishes were filled with 1% agar solution in FSW and left to dry. The 1% agar solution was preheated using a microwave allowing the melting of the agar. Agar dishes were kept in a humid chamber at 4 ^oC until use.

Morpholinos: The MASO against *Sp-Pdx1* was already available in the laboratory where this PhD project was carried out, whereas *Sp-NosA, Sp-NosB,* and *Sp-An* were newly designed morpholinos. Further information and sequences of the MASOs used are shown in Table 2.1.

2.4. Dissociation of embryos/larvae

For the scRNA-seq experiments, sea urchin embryos and larvae were dissociated into single cells (Fig. 2.3) according to the adaptation of several existing protocols (McClay 1986, McClay 2004, Juliano et al. 2014). Specimens were collected at 2, 3, and 5 dpf, concentrated using a 40 μm Nitex mesh filter, and spun down at 500 g for 5 min. Seawater was removed, and larvae were resuspended in Ca^{2+} Mg^{2+} free artificial seawater. Specimens were spun down at 500 g for 5 min and resuspended in dissociation buffer containing 1 M glycine and 0.02 M EDTA in Ca^{2+} Mg^{2+}free artificial seawater. Samples were incubated for 10 min on ice and mixed gently via pipette aspiration every 2 min. From that point and onwards, the progress of dissociation was monitored and fully dissociated cells were obtained in approximately 20 min. Dissociated cells were spun down at 700 g for 5 min and washed several times with Ca^{2+} Mg^{2+} free artificial seawater. Cell viability was assessed using propidium iodide and fluorescein diacetate, and only specimens with cell viability ≥ 90% were further processed. Single cells were counted using a hemocytometer and diluted according to the manufacturer's protocol (10x Genomics). Throughout this procedure, samples were kept at 4 ^oC.

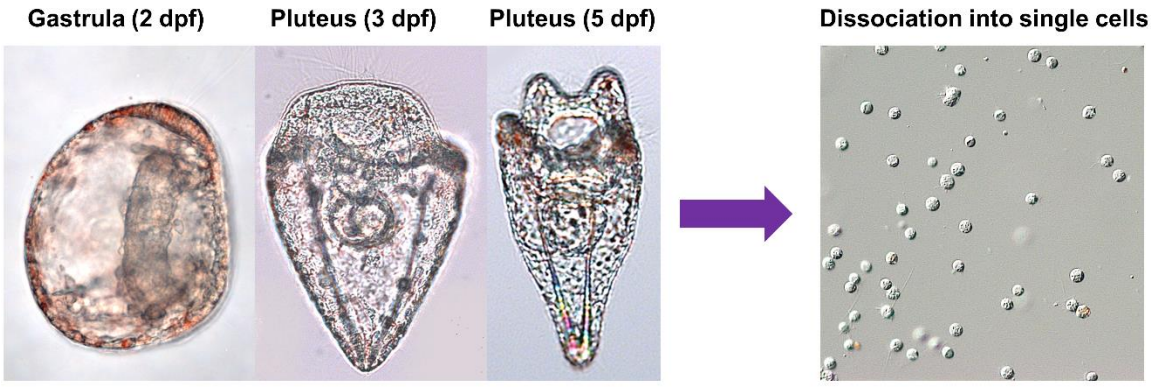


Figure 2.3. Dissociation of embryos and larvae into single cells. DIC images of 2 dpf gastrula, 3 dpf early pluteus, 5 dpf pluteus stages (left), and single cells in suspension (right).

2.5. Single cell RNA sequencing library preparation

ScRNA libraries were constructed using the 10x Genomics single-cell capturing system, and the Chromium Single Cell 3' Reagent kits v2 and v3 were used according to the manufacturer's protocol (10x Genomics). The Chromium 10x Genomics Single Cell Capturing system results in the sequencing of the 3' end of mRNAs coming from individual cells. GemCode system based on a pool of approximately 3,500,000 10x barcodes allows the separate indexing of each cell's transcriptome. The initial step of the process is the generation of GEMs (gel beads in the emulsion). GEMs are generated when Single Cell Gel Beads containing barcoding primer mix [Illumina True Seq Read1, 16 nt 10x barcode, 12 nt unique molecular identifier-UMI and 30 nt poly(dt) sequence], a master mix solution containing cells in suspension and partitioning oil are loaded onto a microfluidics chamber, the Chromium Chip. The Chip is then placed into the 10x Genomics Chromium controller, and GEMs are generated in a process that lasts for 10-17 min. Once GEMs are generated, the gel beads dissolve, and the primers are released and mixed with the cell lysate and a Master mix solution containing reagents for reverse transcription (RT). The sample is collected from the Chip and placed into a thermal cycler according to 10x Genomics guidelines. This results in the production of barcoded complementary DNA (cDNA) coming from poly-adenylated mRNA.

Once cDNA is synthesized (first-strand DNA), silane magnets are used to remove the RT reagents leftovers. The resulting pure barcoded cDNA is amplified via polymerase chain reaction (PCR) to increase cDNA to a sufficient mass for library construction. Following rounds of silane magnet-based purification, enzymatic

fragmentation of the full-length cDNA is applied. P5 and P7 sample index and TrueSeq Read 2 are added to the barcoded cDNA via End Repair, A-tailing, Adaptor Ligation, and PCR. The resulting libraries were sequenced by GeneCore (EMBL, Heidelberg, Germany) for 75 bp paired-end reads (Illumina NextSeq 500).

In all steps, before sequencing, the amount, size, and quality of the cDNA were estimated from the Bioanalyzer trace (Agilent Bioanalyzer High Sensitivity chip).

The 10x Genomics Single Cell capturing system was used to construct *S. purpuratus* single-cell libraries from 3 biological replicates of 2dpf gastrula (sp48_1, sp48_2, sp48_3), 4 biological replicates of 4dpf early pluteus (sp72_1, sp72_2, sp72_3, sp72_4), and 2 biological replicates of the 5dpf pluteus (sp5d_1, sp5d_2). Furthermore, in total 4 technical replicates for four of these biological replicates (sp48_1a, sp48_2a, and sp72_1a, sp72_2a) were produced to which higher sequencing depth was applied.

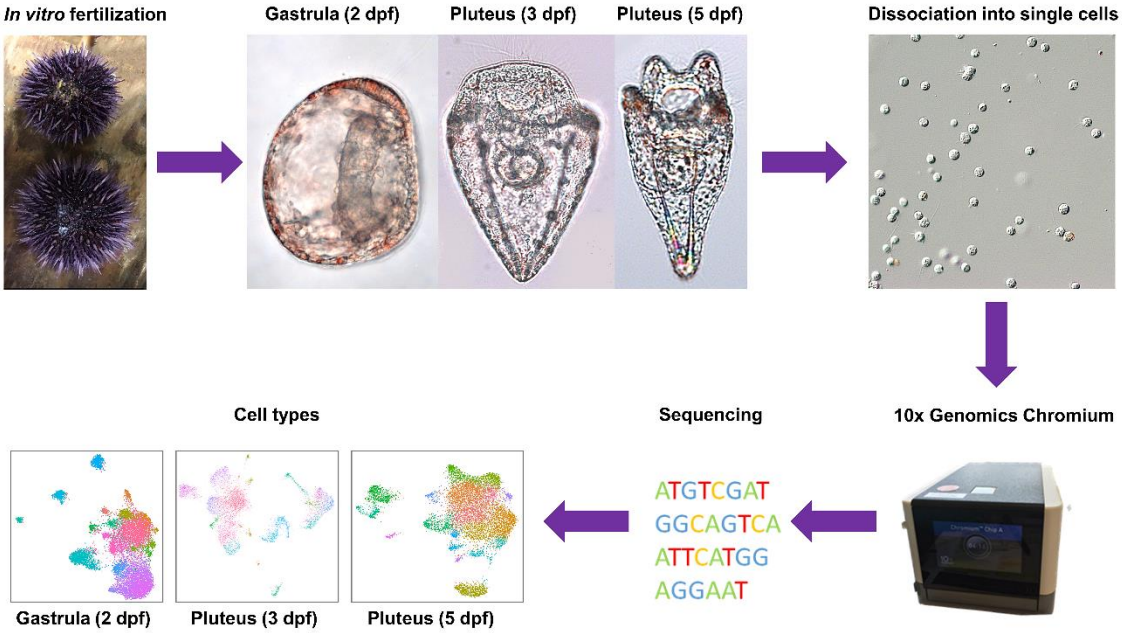


Figure 2.4. From embryo/larva to single cells. Schematic representation of our single cell RNA sequencing pipeline from gamete fertilization to computational analysis.

2.6. Single RNA sequencing analysis

The initial conversion of the BCL read files were converted into FASTQ by Dr. Jacob M. Musser, a postdoc from Dr. Arendt's lab, using Cellranger 3.0.2 (10x Genomics). The single-cell data analysis's initial steps were performed by Dr. Danila Voronov,

a postdoc from Dr. Maria Ina Arnone lab. These steps included the alignment of the single-cell RNA-seq output reads and generation of feature, barcode, and matrices (Cellranger 3.0.2) and the construction of the genomic index in Cell Ranger using the *S. purpuratus* genome version 3.1 (Sea Urchin Genome Sequencing et al. 2006, Kudtarkar and Cameron 2017). *S. purpuratus* Genome v3.1 was chosen as the reference genome due to it having the most complete annotation at the time of analysis and writing. A stringent number of cells per biological/technical replicated was chosen based on Knee plots showing the distribution of barcode counts as well as which barcodes were inferred to be associated with cells, therefore discriminating "cells" from background (Fig. 2.5).

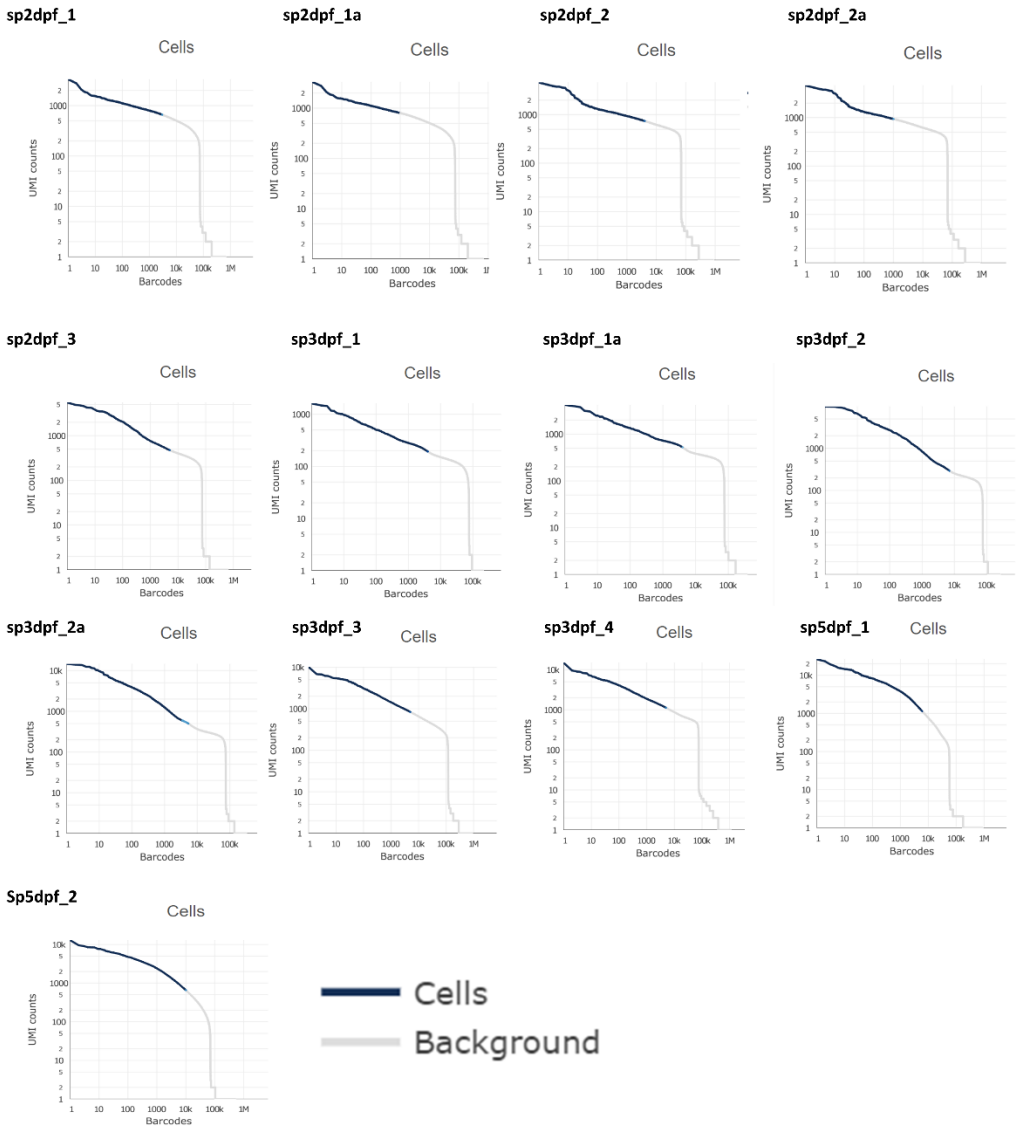


Figure 2.5. Knee plots allowing the discrimination between what is a cell from background noise (non-cells). Cells with high UMI counts (left part of the knee) are considered as real cells.

Cell Ranger output matrices for all biological replicates were used for further analysis in Seurat v3.0.2 R package (Stuart et al. 2019). The downstream analysis in Seurat was performed by me with the assistance of Dr. Voronov. The analysis was performed according to the Seurat scRNA-seq R package documentation (Butler et al. 2018, Stuart et al. 2019). First, the barcodes and genes files were combined so that each transcript's identity can be retrieved. Genes transcribed in less than three cells and cells with less than a minimum of 200 transcribed genes were excluded from the analysis. The cutoff number of transcribed genes was determined based on feature scatter plots and varies depending on the replicate and developmental stage. In detail, based on the violin plots showing the numbers of features and RNA counts (Figs. 2.6, 2.7, 2.8) per biological/technical replicate and developmental stage, the parameters shown in the Table 2.4 were chosen.

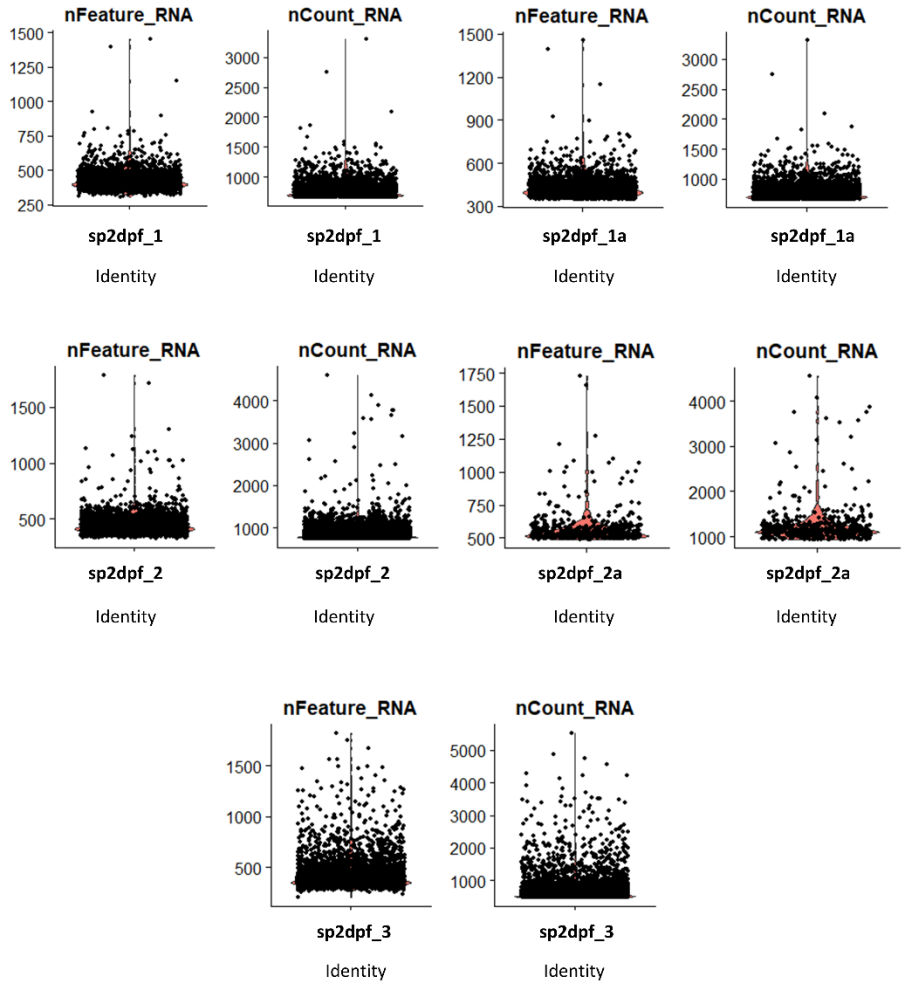


Figure 2.6. Violin plots showing the number of features and RNA counts of the 2 dpf libraries. nFeature_RNA represents the number of genes detected in each cell, while nCount_RNA shows the total number of molecules detected within a cell (UMIs). Low nFeature_RNA indicates either a dead/dying cell or an empty droplet.

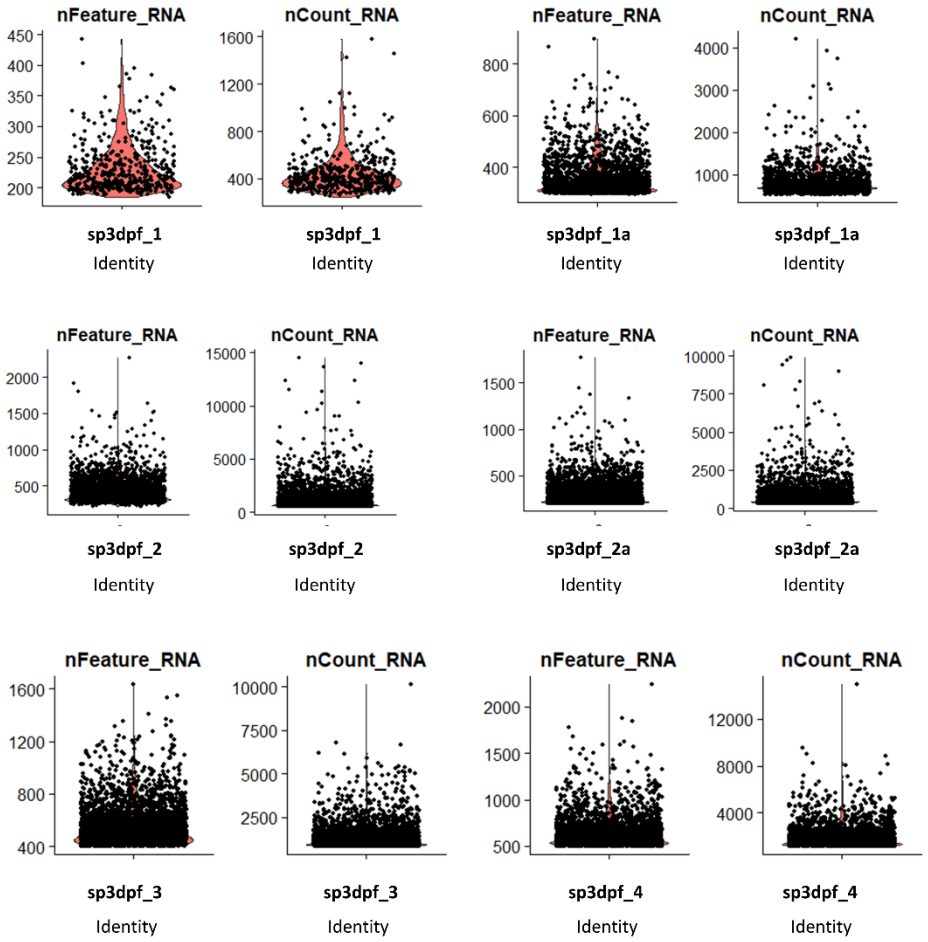


Figure 2.7. Violin plots showing the number of features and RNA counts of the 3 dpf libraries. nFeature_RNA represents the number of genes detected in each cell while nCount_RNA shows the total number of molecules detected within a cell (UMIs). Low nFeature_RNA indicates either a dead/dying cell or an empty droplet.

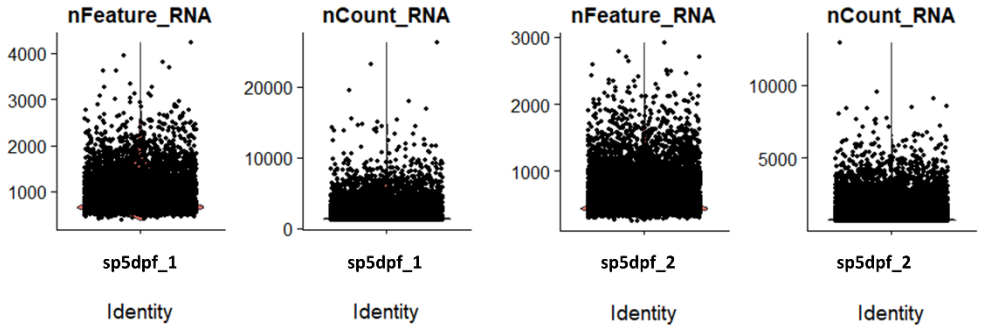


Figure 2.8. Violin plots showing the number of features and RNA counts of the 5 dpf libraries. nFeature_RNA represents the number of genes detected in each cell while nCount_RNA shows the total number of molecules detected within a cell (UMIs). Low nFeature_RNA indicates either a dead/dying cell or an empty droplet.

Datasets were normalized, and variable genes were found using the vst method with a maximum of 2000 variable features. Data integration was performed via identification of anchors between different objects per developmental stage. Next, the datasets were scaled, and principal component (PCA) analysis was performed. The gene variability is extracted and brought out by PCA. The Nearest Neighbor (SNN) graph was computed with 20 dimensions for 2 and 3 dpf datasets and 36 dimensions for the 5 dpf dataset (resolution 1.0) to identify the clusters as suggested by the Elbow plots shown in Figure 2.9. For combined datasets (2 and 3 dpf), SNN graphs were computed with 20 dimensions (Fig. 2.9). Uniform Manifold Approximate and Projection (UMAP) was used to perform clustering dimensionality reduction. Cluster markers were found using the genes detected in at least 0.01 fraction of min. pct cells in the two clusters. Transcripts of all genes per cell type were identified by converting a Seurat DotPlot with all these transcripts as features into a table (ggplot2 3.2.0 R package). Subclustering analysis was performed by selecting a cell type of interest and performing a similar analysis as described above. In detail for the neuronal and anterior neuroectoderm subclustering analyses coming from the 2 dpf datasets, SNN graphs were computed with 3 and 12 dimensions respectively (Fig. 2.9). Regarding the subclustering of the 3 dpf neuronal, apical plate, skeletal and immune cell subclustering analyses, SNN graphs were computed with 11, 13, 5 and 18 dimensions respectively (Fig. 2.9). Finally concerning the subclustering of the apical plate putative broad cell type coming from the merged 2 and 3 dpf datasets, SNN graphs were computed with 20 dimensions (Fig. 2.9). The integration of replicates at different developmental time-points was performed similarly to what has been described above. All resulting tables containing the genes transcribed within different cell types were further annotated, adding PFAM terms (Trapnell et al. 2010, Finn et al. 2014) for associated proteins, gene ontology terms, and descriptions from Echinobase (Kudtarkar and Cameron 2017). Further details about the steps of the computational analysis can be found in the non-book component of the thesis in the files named “Sp2dpf_clustering_analysis. Rmd”, “Sp3dpf_clustering_analysis. Rmd”, “Sp5dpf_clustering_analysis. Rmd”, “Sp2_3dpf_clustering_analysis. Rmd”, “Subclustering analysis. Rmd” R Markdown objects.

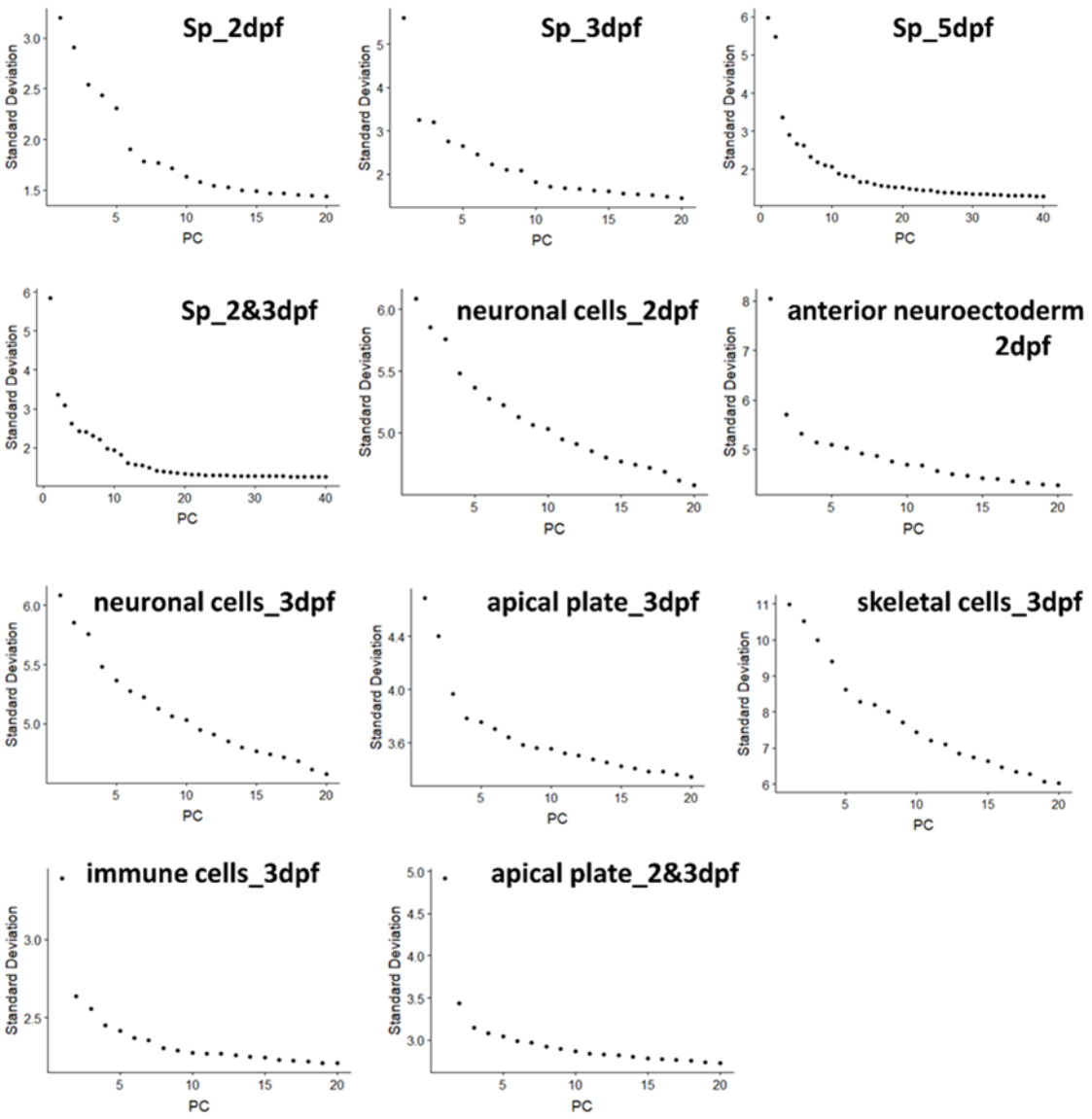


Figure 2.9. Elbow plots showing the standard deviation of the PCAs per dataset. Elbow plots alongside with Jackstraw plots (data not shown) were used to identify the significant dimensions that were included in the subsequent analysis.

2.7. RNA extraction and cDNA synthesis

Total RNA was extracted from 2, 3, and 5 dpf embryos/larvae using the RNAqueous Total RNA Extraction Kit (Thermo Fisher Scientific). A minimum of 500 embryos/larvae were collected in an Eppendorf tube, centrifuged at 700g (4 ^{o}C), FSW was removed and replaced with the kit's lysis solution. Samples were homogenized with vortexing for 1 min, and ethanol was added in a ratio of lysate to ethanol 2:1. The lysate was loaded onto the kit's purification column to which only the nucleic acids are bound, and the rest of the debris passes through the column.

Washing steps were carried out according to manufacturer's instructions. Nucleic acids were eluted and incubated for 30 min with RNAse free DNAse 1 to remove the extracted DNA. The extracted RNA was used as a template for Reverse Transcription (RT) to generate cDNA. First stranded cDNA was synthesized with the SuperScript VILO cDNA synthesis kit according to manufacturers' instructions. The reaction was carried out in a thermal cycler as follows: 25 ^{o}C for 10min, 42 ^{o}C for 60 min, 85 ^{o}C for 5min. The cDNA was stored at -20 ^{o}C until use.

2.8. Embryo/larva fixation

Specimen fixation is one of the most critical steps ensuring the conservation of morphological and intercellular structures. Different fixation protocols were followed to conserve different structures according to the purpose for which the fixed specimens were used. Independently of which protocol was used, the first step of embryo/larva collection is the same. Embryos and larvae of various stages were quickly collected and concentrated from the culture beakers using nylon mesh (e.g., Nitex) with an opening slightly smaller than the embryos.

2.8.1. Fixation for Immunohistochemistry (IHC)

Specimens used for immunohistochemistry were fixed in 4% PFA in FSW for 15 min at RT. Next, the fixation solution was removed, and specimens were washed several times with 1x Phosphate buffered saline in 0.1% Tween (PBST). PBST. Depending on the protein antigen sensitivity, the 1x PBST washing step was followed by incubating the sample in ice-cold methanol for 1 min. Methanol was removed, and specimens were washed several times with 1x PBST. Fixed embryos and larvae were kept at 4 ^{o}C until use.

2.8.2. Fixation for whole-mount Fluorescent *In Situ* Hybridization (FISH)

Concentrated embryos/larvae were placed in Eppendorf tubes, spun down at a low speed (~500g) for 10 min at 4 ^{o}C, and water was discarded. Specimens were resuspended in fixative solution containing 4% paraformaldehyde, 0.1M MOPS pH 7, 0.5M NaCl 0.1%, in nuclease-free water, tubes were gently inverted 2–3 times to

mix the embryos/larvae in the fixative solution uniformly. Specimens were incubated O/N at 4°C. The day after, specimens were washed with 1x MOPS buffer (0.1M MOPS pH7, 0.5M NaCl, 0.1% Tween-20 in nuclease-free water) three times to remove all the fixative solution. Samples were gradually dehydrated with increasing ethanol solutions (30%, 50%, and 70% ethanol/water or ethanol/MOPS buffer solutions). Samples in 70% ethanol can be stored at −20 °C for months.

2.8.3. Fixation for Scanning Electron Microscopy (SEM)

Specimens were fixed in 4% Glutaraldehyde in FSW for a minimum of 1 h at room temperature to O/N at 4 °C. Fixative was removed, and samples were gradually dehydrated with increasing ethanol solutions (30%, 50% and 70%, 80%, 90% ethanol/water) and finally transferred to absolute ethanol. Samples were washed 3 times in absolute ethanol to ensure the removal of all water molecules. Scanning Electron Microscopy preparations are followed, and samples are ready to be imaged.

2.9. Riboprobe synthesis

Antisense riboprobes (RNA probes) are chemically labeled RNA molecules that are complementary to the specific mRNAs targets for which they have been designed against. RNA probes are synthesized by in vitro transcription with DNA-dependent RNA polymerases. Labeling of the synthesized RNA probe can occur during the in vitro transcription reaction by incorporating labeled ribonucleotides to the newly synthesized RNA or post-transcriptionally by using chemicals that label the entire RNA probe molecules. The first step towards probe generation is to design the primers used to clone the gene of interest. Gene cloning is a molecular biology tool used to generate multiple identical copies of a specific DNA piece or gene. In this thesis, two different gene cloning techniques were used to isolate and multiply several genes of interest to produce sufficient mass to *in vitro* synthesize antisense RNA probes.

2.9.1. Gene cloning

One of the gene cloning techniques applied was the traditional DNA cloning of the probe into a vector and used when a probe was needed in high amounts or frequently used while conducting this PhD work. The advantage of this approach is that bacterial stocks containing the plasmid with the insert (gene) of interest can be kept at -80 ºC for years. To this end, the probe sequence was amplified from cDNA isolated from 2, 3, and 5dpf embryos and larvae with PCR, purified with a GenElute PCR Clean-Up purification kit (Sigma-Aldrich), and inserted into a vector (plasmid) containing the Sp6 and T7 RNA polymerases binding sites. In this study, pGEM®-T Easy Vector (Promega) was used as a vector, and the amplified sequence was inserted into the plasmid with the T4 DNA ligation enzyme. T4 DNA Ligase catalyzes the formation of phosphodiester bonds between 5'-phosphoryl and 3'-hydroxyl DNA ends and thus joins both blunt-ended and cohesive-ended DNA fragments. In each ligation reaction, the vector to insert ratio was 1:3 and was performed according to the manufacturer's guidelines (Promega). The DNA ligation reaction was incubated at 15 ºC overnight (O/N), and a fraction of it was used to transform competent bacteria.

Competent bacteria, prepared by the Molecular Biology and Sequencing service at Stazione Zoologica Anton Dohrn, were chemically transformed with the recombinant plasmids. Briefly, 4 µl of the DNA ligation reaction (containing plasmid plus insert) were added into 50 µl of competent bacteria and left to incubate for 40 min on ice. Next, the bacteria-plasmid solution was placed at 42 ºC for 90 s and then transferred and kept on ice for 2-3 min. This process creates a heat shock that leads to bacterial cells incorporating the recombinant plasmids (transformation). Then, the transformed bacteria are placed in 1 ml of preheated at 37 ºC LB broth (NaCl 10 g/l Bacto-tryptone 10 g/l, Bacto-yeast extract 5 g/l) left to incubate at 37 ºC for 1 hour, plated on LB solid medium (NaCl 10g/l Bacto-tryptone 10 g/l, Bacto-yeast extract 5 g/l, agar 15 g/l) in the presence of ampicillin (100 mg/ml) and coated with (100 mM) IPTG and X-Gal (20 mg/ml). Plated bacteria were let to grow O/N at 37 ºC. Whether bacteria were successfully transformed by a recombinant plasmid can be easily observed by the color of their colonies. Competent bacteria originate from a mutated bacterial *E. Coli* strand that cannot produce a functional β-galactosidase enzyme.

On the other hand, the vector contains sequences that code the same enzyme placed out of the translational frame when plasmids are successfully recombined. Therefore, upon provision of the β-galactosidase substrate (X-gal), bacteria that have been transformed by non-recombinant plasmids are able to produce β-galactosidase, catabolize X-gal and form bacterial colonies that are of blue color, whereas the bacteria transformed by recombinant plasmid remain white. This selection method, called blue and white screening allows the identification of the bacterial colonies transformed by a recombinant plasmid and thus with the plasmid containing the gene of interest. Single white colonies are selected and placed into 3 ml of LB broth and ampicillin and let to grow O/N at 37 ^{o}C. A plasmid purification kit (Sigma-Aldrich) was used to isolate the plasmid DNA from the cells, according to manufacturers' instructions.

An alternative, faster method to generate a probe is to design a typical forward primer and a reverse primer that contains at the 5' end the binding sequence for Sp6 (5´-ATTTAGGTGACACTATAG-3´) or T7 (5′-TAATACGACTCACTATAG-3′) RNA polymerase. The advantage of applying this method is that the first PCR product already contains the RNA polymerase binding site and can be used as a template for *in vitro* transcription.

In both approaches, probe length ranged from 500-2000 bp; the probes' size and quality were assessed by running a small portion of each probe on a 1% agarose 1x TAE gel, and their correct identity was verified using sequencing. The online bioinformatics tool Primer 3 (http://bioinfo.ut.ee/primer3/) was used to design primers for gene cloning.

Notes:

- Detailed information on the primers used to clone genes is presented in Table 2.2.
- Probes for *Sp-Pdx1, Sp-Cdx, Sp-ManrC1A, Sp-Mhc, Sp-Rhox3, Sp-Ptf1a Sp-Fgf9/16/20, Sp-Brn1/2/4, Sp-Ngn, Sp-Isl, Sp-NeuroD1, Sp-Soxb2, Sp-An, Sp-Trh, and Sp-Salmfap* were produced as previously published [*Sp-Pdx1, Sp-Cdx* (Cole et al. 2009), *Sp-ManrC1A* (Annunziata et al. 2014), *Sp-Mhc (Andrikou et al. 2013), Sp-Rhox3(Perillo 2013), Sp-Ptf1a(Perillo et al. 2016)*, *Sp-Fgf9/16/20* (Andrikou et al. 2015), *Sp-Brn1/2/4* (Cole and Arnone 2009),

Sp-Ngn, Sp-Isl, Sp-NeuroD1 (Perillo et al. 2018), *Sp-SoxB2* (Anishchenko et al. 2018), *Sp-An, Sp-Trh, Sp-Salmfap* (Wood et al. 2018).

2.9.2. Labelled Riboprobe synthesis

As previously mentioned, probe labeling can either be carried out during the *in vitro* transcription reaction or post-transcriptionally. In the current PhD work, both approaches have been applied to synthesize different types of probes.

Digoxigenin (Roche) or Fluorescein (Roche) labeled ribonucleotides were used to generate digoxigenin or fluorescein-labeled RNA probes during *in vitro* transcription. 500–1,000 ng of linear DNA were used as a template for the *in vitro* transcription reaction, and the reaction was incubated at 37 ^{o}C for 2 hours. Then RNase free DNase I was added (1 U/µl) to remove the DNA template, and the mix was incubated for another 20 minutes at 37 ^{o}C. Unincorporated nucleotides and the reagent leftovers were removed by purifying the RNA probes using the Mini Quick Spin RNA Columns G-50 Sephadex (Roche).

Dinitrophenol (DNP) labeled probes were labeled post-transcriptionally. The first step is to in vitro transcribe unlabeled RNA starting from 500–1,000 ng of linear DNA; the reaction was carried out at 37 ^{o}C for 2 hours according to the manufacturers' instructions and purified using the Mini Quick Spin RNA Columns G-50 Sephadex (Roche).

2.5 – 5µg of pure unlabeled RNA is used as a template for the DNP Labelling kit (Mirus). The DNP labeling kit consists of the DNP label, a linker that facilitates electrostatic interactions with nucleic acids, and the reactive alkylating group that attaches the reagent to non-base-pairing sites on the guanine ring. The labeling reaction was carried out at 37 ^{o}C for 2 hours, and the labeled RNA probe was purified as described above.

RNA probes were stored at −80 ^{o}C until use.

2.10. Whole-mount Fluorescent *In Situ* Hybridization (FISH)

FISH was performed using the Tyramide Signal Amplification (TSA) Kit (Perkin Elmer). TSA reagents allow the fluorescent detection of single mRNA molecules. Fluorescence is produced once the HRP enzyme conjugated to the antibodies that recognize the chemicals with which the probe is labeled (anti-DIG, anti-DNP, anti-FLUO. HRP transforms the TSA reagents to free radicals that form bonds with tyrosine residues proximal to HRP. The TSA reagents used in this thesis are cyanine 3 (cy3, excitation 550 nm, emission 570 nm) and cyanine 5 (cy5, excitation 648 nm, emission 667 nm). The protocol used for single or double FISH (Perillo et al. 2021) is as follows, and a summary of the steps is presented in Figure 2.10 :

1. Embryos or larvae stored in 70% ethanol were gradually rehydrated with decreasing ethanol concentration in MOPS buffer and then washed x3 with MOPS buffer.
2. Specimens were washed x2 in pre-warmed at 50 ^{o}C hybridization buffer (70% formamide, 100 mM MOPS pH 7, 500 mM NaCl, 0.1% Tween 20, 1 mg/ml BSA) and left in a 50 ^{o}C incubator for 3 hours to pre-hybridize.
3. The DIG-, FLUO- and/or DNP-probes were added at a final concentration of 0.1–0.5 ng/µl in hybridization buffer. The final volume of the hybridization was at least ten times the embryo/larva volume. The tube with the probe dilution was heated at 65–70 ^{o}C for 10 minutes, spun briefly to get all the probe solution to the bottom of the tube, and left for 5 min on ice. Then the pre-hybridization buffer was removed, and the diluted probe containing solution was added to the samples.
4. Specimens were left to hybridize for 5-7 days at 50 ^{o}C. During this time, the RNA probe binds the endogenous RNA of interest.
5. After hybridization, samples were washed with hybridization buffer for at least 3 h at 50 ^{o}C, replacing the buffer twice with pre-warmed (50 ^{o}C) buffer.
6. Then the embryos/larvae were washed x3 with MOPS buffer at room temperature.
7. Specimens were incubated in the PerkinElmer blocking reagent that comes with the kit for 30 minutes at room temperature. Then the blocking solution was removed, replaced with a dilution of 1: 1,000 of either anti-DIG or anti-DNP or anti-FLUO horseradish peroxidase (HRP) Fab fragments (Roche)

in the PerkinElmer blocking buffer and incubated at 37 ^{o}C for 1 hour or at 4 ^{o}C overnight.

8. The excess of the unbound antibody was removed by washing with MOPS buffer x5 at room temperature.
9. The signal was detected by staining for 15–30 minutes in 1:400 cy3 in Perkin Elmer amplification diluent.
10. Unbound cy3 was removed by washing the samples in MOPS buffer x5.
11. In the case of single FISH 1:10,000 DAPI (from a 10mg/ml stock prepared in water) in MOPS buffer was added to image nuclei.
12. In the case of double FISH, the HRP conjugate to the antibody detecting the first probed was inactivated by washing the samples x2 with 1% H_2O_2 in MOPS buffer for 30 min.
13. Specimens were washed in MOPS buffer x5 to eliminate all the H_2O_2.
14. Samples were incubated again in Perkin Elmer blocking reagent for 30 min.
15. Embryos/larvae were incubated with 1:1,000 of the appropriate antibody depending on how the second RNA probe was labeled (anti-DNP-HRP, anti-FLUO-HRP antibodies in blocking buffer for 1 hour at 37 ^{o}C or overnight at 4°C.
16. Specimens were washed 5x in MOPS buffer to remove excess antibody.
17. Samples were washed in Perkin Elmer amplification diluent x2.
18. The signal was detected by staining for 15–30 minutes in 1:400 cy5 (from the TSA kit) in amplification diluent.
19. Specimens were washed x5 in MOPS buffer.
20. 1:10,000 DAPI (from a 10 mg/ml stock prepared in water) in MOPS buffer was added to image nuclei.
21. Specimens were mounted for imaging with a Zeiss LSM 700 confocal microscope.

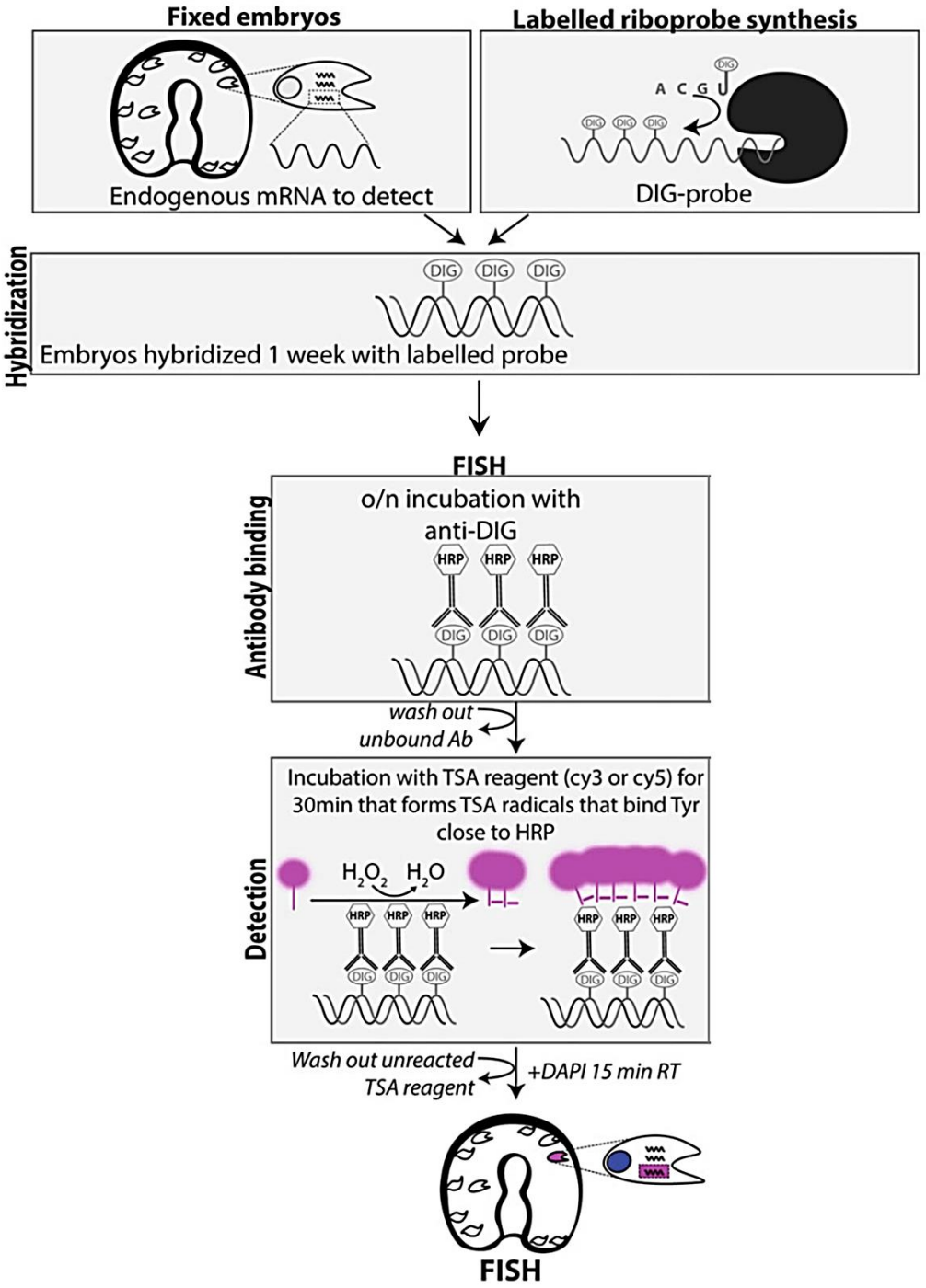


Figure 2.10. Fluorescent *in situ* hybridization. Schematic representation summarizing the FISH technique as described above. (Figure adapted from Perillo et al. 2021).

2.11. Immunohistochemistry (IHC)

Embryos/larvae fixed either for IHC or in FISH were used for immunohistochemical detection of protein antigens. In case fixed specimens for FISH were used, samples

were gradually rehydrated with decreasing ethanol concentration in 1x PBST and then washed x3 with 1x PBST. The IHC protocol was carried out as follows:

1. 1x PBST was removed, and specimens were incubated with blocking solution (1% BSA, 4% Sheep serum diluted in 1x PBST) for 1h at room temperature or overnight at 4 ^{o}C.
2. The primary antibody(ies) of interest was (were) diluted in the appropriate concentration in blocking solution. The blocking solution was removed from the specimens and replaced with a blocking solution containing the diluted antibody(ies). Samples were incubated either for 1 hour and 30 min at 37 ^{o}C or overnight at 4 ^{o}C. For further information on the specific characteristics and dilutions of each antibody use, please refer to Table 2.3.
3. Specimens were washed x5 in 1x PBST to remove excess antibody.
4. The appropriate secondary Alexa Fluor conjugated antibody(ies) (according to the organism in which the primary antibody has been raised) was (were) diluted 1:1,000 in 1x PBST. 1x PBST was removed, and samples were incubated for 1 h at room temperature.
5. Specimens were washed x5 in PBST to remove excess antibody(ies).
6. 1:10,000 DAPI (from a 10 mg/ml stock prepared in water) in 1x PBST was added to image nuclei.
7. Specimens were mounted for imaging with a Zeiss LSM 700 confocal microscope.

2.12. Combination of FISH and IHC

FISH and double FISH experiments can be combined with indirect immunofluorescence enabling the user to further characterize those cells by adding information provided by the fluorescent signal of a specific antibody to a protein of interest. In this PhD work, immunohistochemistry was combined with FISH after FISH was completed.

1. At the end of the FISH protocol, MOPS Buffer was removed specimens were resuspended 1x PBST.
2. Samples were washed 4 times with 1x PBS to ensure that all MOPS buffer residues were removed.

3. 1x PBST was removed and specimens were incubated with blocking solution (1% BSA, 4% Sheep serum diluted in 1x PBST) for 1h at room temperature or overnight at 4 ^{o}C.
4. The primary antibody of interest was diluted in the appropriate concentration in blocking solution. Blocking solution was removed from the specimens and replaced with blocking solution containing the diluted antibody. Samples were incubated either for 1 hour and 30 min at 37 ^{o}C or overnight at 4 ^{o}C.
5. Specimens were washed x5 in 1x PBST to remove excess antibody.
6. The appropriate secondary Alexa Fluor conjugated antibody (according to the organism in which the primary antibody has been raised) was diluted 1:1,000 in 1x PBST. 1x PBST was removed, and samples were incubated for 1 h at room temperature.
7. Specimens were washed x5 in PBST to remove excess antibody.
8. 1:10,000 DAPI (from a 10 mg/ml stock prepared in water) in 1x PBST was added to image nuclei.
9. Specimens were mounted for imaging with a Zeiss LSM 700 confocal microscope.

2.13. EdU labeling

To investigate the spatial distribution of proliferating cells, cell proliferation assays were carried out using Click-It EdU Cell Proliferation Kit for Imaging Alexa FlourTM 647 (Thermo Fisher Scientific). Larvae were treated with EdU at a final concentration of 10 μM in FSW and let to develop for 2 hours. Samples were fixed in 4% PFA in FSW for 15 min (RT) and washed several times with 1x PBST. PBST was removed, replaced by 100% methanol for 1 min (RT), and followed by several washes with PBST. After this step, one can continue with either developing the EdU signal or pairing this assay with immunohistochemical detection of a protein of interest, as described above. To develop the EdU signal, the Click-iT TM reaction mix was prepared according to the manufacturer's guidelines. PBST was removed, and the reaction mixture was added to the samples for 30 min (RT). Larvae were washed several times with 1x PBST, and 1: 10,000 DAPI in 1x PBST was added to image nuclei.

2.14. TRIM treatments in sea urchin larvae

In order to understand the role of Nitric Oxide (NO) during sea urchin development, we used the 1-(α,α,α-trifluoro-o-tolyl)-Imidazole (TRIM, Cayman Chemical, stock solution in DMSO) that inhibits proper NOS (Nitric oxide synthase) activity by interfering with the binding of both L-arginine and tetrahydrobiopterin to their respective sites on NOS enzymes and thus limiting the production of NO. A range of inhibitor concentration and timing were tested. When the inhibitor was added before gastrulation was complete, it resulted in developmentally arrested embryos. To avoid such possible off-target effects, the inhibitor was applied to already gastrulated embryos (at 3 dpf). The *in vivo* experiments were performed at 500, and 100 µM of TRIM diluted in FSW. Treatments were carried out in 6 well plates at 15 ^{o}C. An equal number of 2 dpf embryos were placed into two wells containing TRIM and DMSO (control), respectively. Larval morphology was assessed at 3 and 4 dpf using a Zeiss Apotome 2 microscope. Larvae were fixed for FISH, IHC, and SEM following the protocols described above.

2.15. DAF-FM diacetate assay

NO detection was performed with 4-amino-5-methylamino-2',7'-difluorofluorescein diacetate (DAF-FM-DA). DAM-FM is a cell-permeable and non-fluorescent reagent that, after reaction with NO, forms the fluorescent compound benzotriazole (Kojima et al. 1999). Wild type, control, and Trim treated larvae were incubated for 20 min in the dark with 5 µM DAF-FM-DA in FSW. Next, the specimens were washed several times in FSW and observed under a Zeiss Apotome 2 microscope.

2.16. Scanning Electron Microscopy (SEM)

Fixed samples (as described above) were subjected to critical point drying, were coated with gold or platinum in a Leica EM ACE 200 Sputter Coater, and observed under a JEOL JSM-6700F scanning electron microscope.

2.17. Gene annotation and orthology assessment.

Gene annotation and orthology assessment has been performed by Echinobase (www.new.echinobase.org) and is based on the DRSC Integrative Ortholog Prediction Tool (DIOPT), which collects the output from several algorithms. For each sequence in Echinobase, orthology predictions are presented for multiple species with direct links to other MODs (e.g., ENSEMBL). Moreover, a gene is a duplicate if more than two SPU genes have an exact start and end coordinates with a single WHL ID, while in the specific cases of *Sp-Pax6*, *Sp-NosB* genes and *Sp-An*, two WHL IDs correspond to the same gene and that the reason why the gene names appear twice in the relative plots.

2.18. Diagrams, graphs, and figures

Figures, diagrams, and graphs were made using Microsoft PowerPoint and Excel Professional Plus 2016, Photoshop CC 2015, R Studio v1.2.1502, and Image J v1.53c. Gene Regulatory Networks were drawn using BioTapestry (Longabaugh 2012).

TABLES

Table 2.1. Sequences of the morpholinos used in this study.

Gene name	Sequence	Type	Reference
Sp-Pdx1	AGTACGCGGGATTGTTCCCTTCCAT	Translation Morpholino	Cole et al. (2009)
Sp-An	TAGATATGCGTTTCGTGACATC	Translation Morpholino	unpublished
Sp-NosA	AATTCGCTCAGAGTTCGGAAGGCAT	Translation Morpholino	unpublished
Sp-NosB	GTCCTGATCCATTTTAAATCACTTA	Translation Morpholino	unpublished

Table 2.2. Primers used to clone genes.

Gene name	WHL ID	Forward Primer	Reverse Primer
Sp-Ahrl	WHL22.256860	GCTTTCATCACGTCGACCTC	TGGCTGAAAACACTCCATGC
Sp-Bra	WHL22.600041	CACCCTTTCACCGCCACTAT	AGAGATCATGTCCGTCGTTTCTA
Sp-Chrna9_4	WHL22.694407	ACACGGAAAACCGTCTCATC	CCGACAGTGCCTCTTTCTTC
Sp-Cyp2L42	WHL22.610666	TGACAGCTCGAAGAAGCTCA	GGACCAGGAAAGTGGTGAAA
Sp-Ddc	WHL22.176642	CCACCGATGACAAAGGTTCT	CCAGGCATACGTCATGTGTC
Sp-Delta	WHL22.423696	GTGCGCTGAAACCTACTACG	CCTCAGTTGCGTGTAATCCG
Sp-Emx	WHL22.113468	CGGCGGAGTTATTTCGACC	TCCCTCTCCCTCTCTTCGAT
Sp-FbsL_2	WHL22.397681	TTTCACACGCTTCATCATCC	TGATAGCTGCTGCTCCACAT
Sp-Fcoll/II/III	WHL22.541149	GCAAGTGTTCGCTTTGATGA	TGTCCGTCACAATGACACCT
Sp-Fgfr1	WHL22.323968	AGATGGCGTTCAGAAGTGGA	CATAGCCAGCTCCGAGATCA
Sp-FoxA	WHL22.439762	TCCCACCCCAACCGACTCCG	CGTCCCTTCGAAATGAATGGACAGGG
Sp-FoxABL	WHL22.615153	TGATAAGATACCTATGTCTTATGAACG	TCGATTAAAATGAAGGTAATGTGC
Sp-FoxI	WHL22.535569	ATCAGACAGGCAACATCCAG	TGTCATGACCGAGCAGACTC
Sp-Foxq2	-	TTGAAAACCTTGCCCAGAGT	TGCATCGCTGGTGGTAGTAG
Sp-Frizz5/8	WHL22.42488	GTGGAACAATCCATCAGTTG	CGTGGTTGCCTACGTAACAG
Sp-Gsc	WHL22.531818	CTCTCCATGAATGCCGCATC	TCATCGTCCGCAGAGTACTC
Sp-Hbn	WHL22.523959	CTCGAGTCCCAGTGGGTTTA	ACTGAGAGCGGTGAACTTGG
Sp-Hnf4	WHL22.35553	GGTGGACAAGGACAAGAGGA	TGAGCTGTAGCAGCTGAGGA
Sp-Hypp_1249	WHL22.480550	AAGCCTACCGAGTCAGATTGC	TTACTGCTTGAAAATAAACATTGAGA
Sp-Hypp_2386	WHL22.239326	TGTCCTCGCTCAGGAAGATT	CGGCTTTATAGGCCAATCC
Sp-IrxA	WHL22.651130	TGCGTATGCTGACTTGAAGG	GGCGGTGATACCACTCTGAT
Sp-Kp	WHL22.176298	TCACGGTCCTATTGGCTCTC	GAAAGGCAGGAAGTGGGGTA
Sp-MacpfA2	WHL22.302097	TACCCCAGTCATCCTTCTCG	GTTCCAGGTGAACTCCGTGT
Sp-Mlckb	WHL22.24095	TGCACCAAGGATACCATCAA	ACAGAGCAGCACCAGTTCCT
Sp-MsxI	WHL22.404908	CGCCACTCACCAACTATTCG	ATGGGAAGGAGAAGCAGCAT
Sp-Nacha6	WHL22.694414	ATGGAGGCCAGACATTGTTC	GCAAGAAGGAACACCCACAT
Sp-Nk1	WHL22.152063	CCCACCATCCGAGTAACATC	GACCATGCATGTGCGTAAAC
Sp-Nkx6.1	WHL22.567494	TCCGAATCGATGATTTCTCTG	TCGTCAATGTCTGCTTCGAC
Sp-NosA	WHL22.55699	CCTGAAGGGTACCGCACTCG	CACCCATACCGTCTGCTGTCC
Sp-NosB	WHL22.504131 WHL22.504131	GATATATCGTGTACAGACTCAC	GTTTTGGTGAGCCAAGCATCG
Sp-Opsin3.2	WHL22.338995	CGCCCTCTACCTGACCTTAG	GCGCGAAAACTCTGCTGATA

Sp-Otp	WHL22.286934	CACCGATGAACGACTCCTTT	ATCAATGTCGCACCAATTCA
Sp-Otx	WHL22.532435	CGATCATTTAGCAGCGAACA	AGAGCTGCGTTCAAGGTCAT
Sp-Pax6	WHL22.585629 WHL22.585512	ATTGGTGAGATGGGTGGGTA	GCTGAAGATGACGAGGATGC
Sp-Prox1	WHL22.531966	GGGATCCTATCGCATCTTCA	CTCCGGTACCGTGAACTCAT
Sp-Rfxc1l	WHL22.741295	CGTGATCAAGTTGGTGGTTG	TCCGAATCAGATGGAAGGAC
Sp-Serp2; Sp-Serp3	WHL22.204515	TACGAAAATGCGTGGATTGA	GGATCCGAGACATCCAAGAA
Sp-Six3	WHL22.121654	ATGGCTTGCGATATGAGCGA	TCTGGCGCGTAGAACACTTT
Sp-SoxC	WHL22.622787	GCAGTCACAGCAGATTCCAA	AATCATGGACGAAAGCAGCC
Sp-Spec2a	WHL22.416148	GCGATGGCAAAATCACTTCT	ATCCGTTGTGATCCTTGTCC
Sp-Tbx2/3	WHL22.457020	TTCTTTCAAAGTCCGCGTCT	AAAGAAGTCCTGCAGCGTGT
Sp-Tph	WHL22.635790	ATCGAATCGAGAAAGGCTCA	GGTCAATTCGTCTCGGACAT
SPU_002797	WHL22.296833	ACAACCGTCTCCATGACTCC	ACTCCTGCCTCTCCTTCCTC
SPU_008104	WHL22.637579	TCTTCTCCTTCTCGGCTTTG	TGAAGTTTTGCACGTTGAGG
SPU_016308	WHL22.79741	CGTCTGTGTACATGGTGTGC	GTGTTGCTCTTGCTCCCAAT

Table 2.3. Antibodies used in this thesis.

Gene name	Type	Dilution
Tyrosine hydroxylase (TH)	polyclonal anti-rabbit	1:50
Choline acetyltransferase (Chat)	polyclonal anti-rabbit	1:100
Serotonin	polyclonal anti-rabbit	1:1,000
uNOS	polyclonal anti-rabbit	1:100
Acetylated a-tubulin	monoclonal anti-mouse	1:200
Beta-tubulin	monoclonal anti-mouse	1:100
Synaptotagmin B	monoclonal mouse	undiluted
Msp130	monoclonal anti-mouse	undiluted
Pdx1	polyclonal anti-rabbit	1:500
Nkx2.1	polyclonal anti-rabbit	1:600
Soxb2	monoclonal anti-mouse	1:400
Endo1	monoclonal anti-mouse	1.1
An	polyclonal anti-rabbit	1:200

Table 2.4. Parameters used in the initial steps of the clustering analysis.

Replicate/developmental stage	minimum number of cells	minimum number of features
sp2dpf_1	3	300
sp2dpf_1a	3	350
sp2dpf_2	3	300
sp2dpf_2a	3	500
sp2dpf_3	3	200
sp3dpf_1	3	200
sp3dpf_1a	3	300
sp3dpf_2	3	200

sp3dpf_2a	3	200
sp3dpf_3	3	400
sp3dpf_4	3	500
sp5dpf_1	3	250
sp5dpf_2	3	250

Contribution Statement

Dr. Maria Ina Arnone has performed the majority of microinjection experiments. The author of this thesis has performed a couple of *Pdx1* morpholino microinjection procedures, has prepared the necessary solutions, for example, PABA-FSW, microinjection solutions, and equipment such as protamine and agar coated plates, as well as microinjection needles.

Dr. Danila Voronov, a member of the Arnone group at the time of the collaboration, has contributed to analyzing the scRNA-seq data. Dr. Jacob Musser, a postdoc from Dr. Detlev Arendt lab, assisted in constructing the first libraries and performed the initial conversion of the BCL read files into FASTQ while Drs. Vladimir Benes and Bianka Baying (GeneCore, Heidelberg, Germany) performed the sequencing of the single cell libraries. The author of this PhD work has optimized the dissociation protocol, constructed the single cell libraries, performed the R studio-based scRNA-seq data analysis, cloned genes, and performed fluorescent *in situ* hybridizations, immunostainings, and combination of the two on scRNA-seq predicted cluster markers. Moreover, the author of this thesis has identified and characterized the different computationally produced cell types and performed the rest of the computational analyses.

Dr. Giovanna Benvenuto at Stazione Zoologica Anton Dohrn has contributed to the imaging process and development of the fixation protocol used for SEM and sample observation. The author of this thesis has equally contributed to all these processes.

Drs. Filomena Caccavale and Salvatore D'Aniello, contributed to the Nitric Oxide related part of the thesis. They were involved in the computational recognition, cloning, and synthesis of the riboprobes for the two sea urchin *nitric oxide synthase* (*NOS*) genes. Dr. Filomena Caccavale and the PhD student Maria Cocurullo designed the morpholino antisense oligonucleotides targeting the two *Nos* genes. The author of this thesis contributed to synthesis of the probe and performed the experiments to characterize their domains of expression and function.

Maria Cocurullo also assisted in cloning gene markers and synthesize antisense probes for several gene markers. Inés Fournon Berodia also assisted in cloning and preparation of probes.

Dr. Margherita Perillo contributed to developing the FISH protocol and the initial characterization of pre-pancreatic neuronal types by cloning pre-pancreatic genes of interest and performing several *in situ* hybridizations and immunostainings. The author of this thesis performed *in situ* hybridizations and immunostainings on control and Pdx1 knockdown embryos/larvae and a thorough analysis of the molecular fingerprint of the *Pdx1/Brn1/2/4* double-positive pre-pancreatic neurons.

All other work presented in this thesis was performed by the author of the thesis.

Chapter 3

Embryonic and larval cell types at a single cell resolution

One of this thesis's main goals is to assess the cell type complexity at a molecular level during *S. purpuratus* embryonic and larval development. To this end, I performed scRNA-seq on three key developmental stages, and this chapter contains the results of the subsequent scRNA-seq analysis. By the end of this analysis, I was able to identify distinct cell types, novel genes markers and provide a detailed overview of the transcriptomic signature and gene regulatory wiring of the early pluteus stage.

3.1. Identification of the cell types in place during sea urchin development

The first aim of my PhD thesis is to investigate at a molecular level the composition and identity of the cell types in place at 2 dpf gastrula, 3 dpf early pluteus, and 5 dpf pluteus developmental stages. Until now, most of our current knowledge on the identity of sea urchin cell types comes from traditional approaches such as identification of molecular markers, perturbation of gene expression, and fate mapping. These approaches contributed to drafting several close to completion GRNs for several cell types at multiple stages of sea urchin early embryonic development.

Recent technological advances in microfluidics and nucleic acid barcoding allow the high-throughput recognition of an organism's cell types at a single cell level, though scRNA-seq. This method has been successfully applied to a variety of different taxa spanning from insects (Davie et al. 2018, Severo et al. 2018, Cho et al. 2020), cnidarians (Sebe-Pedros et al. 2018), and sea squirts (Cao et al. 2019, Sharma et al. 2019) to more complex ones such as zebrafish (Wagner et al. 2018, Chestnut et al. 2020) mice (Nestorowa et al. 2016, Jung et al. 2019, Ximerakis et al. 2019, Yu et al. 2019, Qi et al. 2020) and even humans (Yu et al. 2019, Esaulova et al. 2020, Qi et al. 2020, Zhao et al. 2020), revealing the cell types and their distinct transcriptomic signatures.

To achieve the first goal of this PhD project, I set up a scRNA-seq protocol for the sea urchin embryo and larva, and this chapter contains the results of the scRNA-seq analysis performed on single-cell libraries originating from the aforementioned developmental stages. These results provide a more in-depth understanding of the molecular fingerprint and identity of the sea urchin *S. purpuratus* cell types present in those developmental stages. Due to the current lack of information on the molecular diversity and regulatory dynamics active at the 3 dpf early pluteus stage and the fact that this developmental time-point marks the transition from embryonic to larval lifestyle, 3 dpf pluteus stage was chosen for further characterization and regulatory analysis.

3.1.1. From embryo/larva to single cells

The first step towards unraveling the cell type composition of an organism at a single cell level, is to dissociate the specimen into single cells. To achieve this goal, a gentle enzyme-free dissociation protocol that produces a high number of intact single cells was developed. The dissociation protocol used in this thesis is based on the suspension of embryos and larvae in calcium/magnesium-free artificial seawater (Ca^{2+} Mg^{2+} free artificial seawater). The dissociation is based on the reduction of extensive cell contacts due to the calcium and magnesium absence, allowing the isolation of single cells after applying a gentle mechanical force (pipette aspiration). Glycine and EDTA's presence in the dissociation medium ensured the decompaction of cells and the absence of calcium anions and thus aggregation events.

3.1.2. Generation of cell type atlases

Embryonic and larval cultures coming from 3 biological replicates of 2 dpf gastrula, 4 biological replicates of 3 dpf early pluteus, and 2 biological replicates of the 5 dpf pluteus developmental stages were used for single cell RNA sequencing analysis. Furthermore, in total 4 technical replicates for four of these biological replicates (two for gastrula and two for pluteus stages) were produced and incorporated into the analysis to which higher sequencing depth was applied (Fig. 3.1 A). The capture of the single cells and the cDNA library preparation was done using the 10x Chromium scRNA-seq technology, and the resulting libraries were computationally integrated and further analyzed using Seurat (Butler et al. 2018, Stuart et al. 2019). Integration was performed both on libraries coming from the same developmental time-point and across time-points. The integration success was determined by producing individual for each replicate UMAPs and assessing their overlap (Fig. 3.1 B, C, D). In detail, the integration of libraries coming from the same developmental time-point was successful, since all the produced clusters consisted of intermingled populations of cells originating from all replicates (Fig. 3.1 B, C, D). Regarding the cross-stages data integration, although feasible (data not shown), only the integration of the 2 and 3 dpf stages (Fig. 3.15) is reported in this thesis. The reasoning for this, relies on the low resolution of the 5 dpf datasets (see below), that compromises the clustering analysis of all integrated datasets.

In total, single cell transcriptomes from 15,341 gastrula, 19,699 early pluteus, and 12,097 late pluteus cells were included in the final analysis. Uniform Manifold Approximation and Projection (UMAP) visualization of the integrated datasets revealed 20 putative broad cell types for the 2 dpf gastrula stage, 21 for 3 dpf, and 20 for 5dpf plutei stages.

The term putative broad cell type was chosen because the scRNA-seq computationally produced clusters often do not correspond to 1:1 cell types *in vivo* and may contain different subtypes of the same cell type or different cell states (Shekhar and Menon 2019).

The number of the putative broad cell types produced is directly linked to the analysis resolution chosen (in this case, the stringent resolution of 1.0 was chosen). Each putative broad cell type can be further sub-divided into several sub-types. Since with the resolution 1.0, most of the cell types expected to be present at these developmental stages were revealed, further characterization of the cell types was performed without changing the resolution. When needed to characterize a putative broad cell type further, sub-clustering and re-analysis of the cell type of interest was applied.

Several studies aiming to decipher the *S. purpuratus* embryo and larva cell type diversity have been carried out through the years. Such studies span from lineage tracing experiments (Angerer and Davidson, 1984, Cameron et al., 1987) to studies aiming to thoroughly characterize specialized cell types. For instance studies carried out in *S. purpuratus* embryos and larvae, have demonstrated that the skeleton of the early pluteus larva is produced by an average of 32 skeletogenic cells (Rafiq et al. 2012, Rafiq et al. 2014).

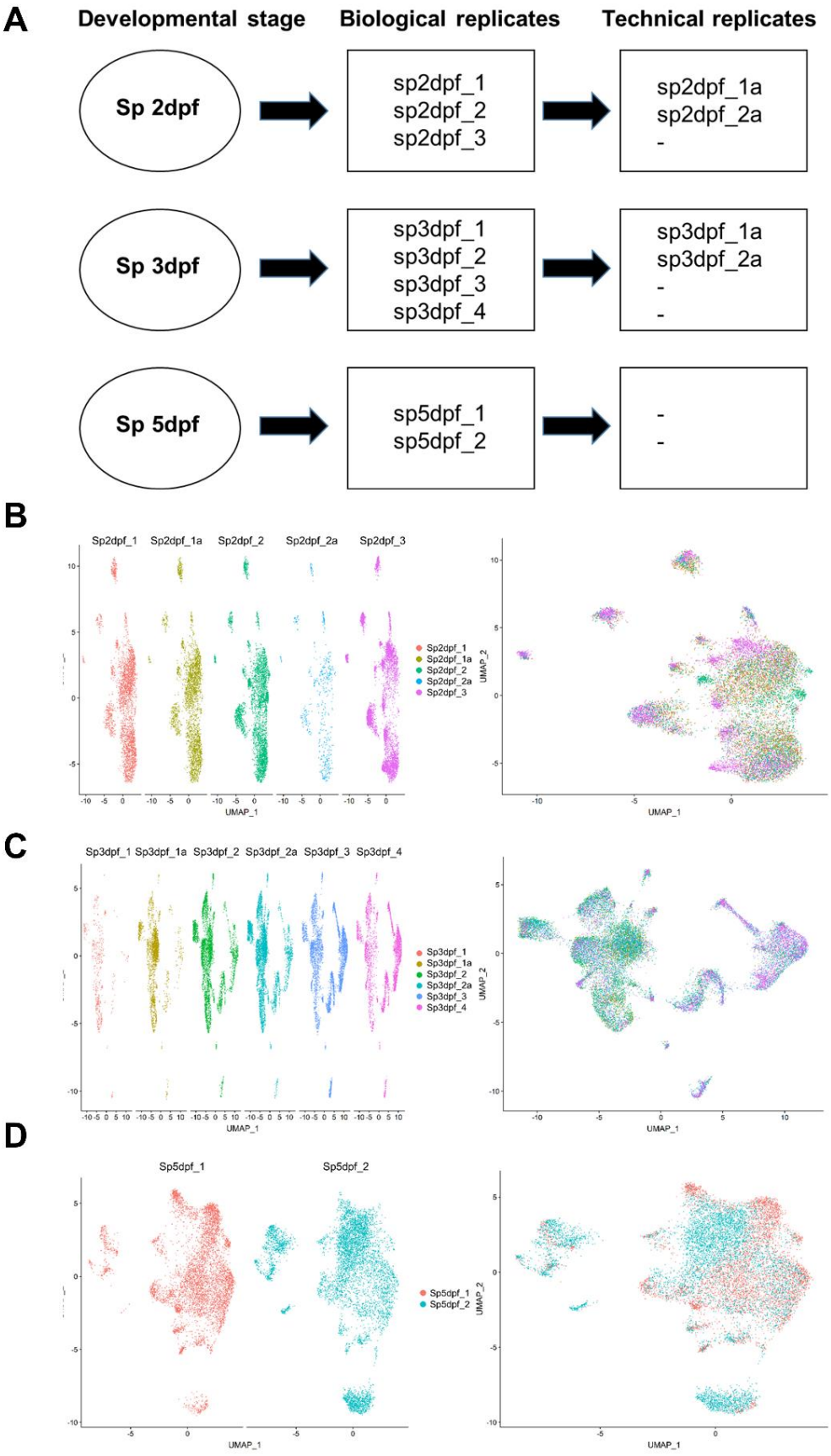


Figure 3.1. Replicate integration. A) Diagram showing the number of samples used per developmental time-point/replicate. UMAPs split and overlay by original identity showing the overlap of the different biological and technical replicates for 2 dpf (B), 3 dpf (C), and 5 dpf (D) developmental stages.

Using such examples and extracting the number of cells constituting the different computationally identified cell types, it is evident that the number of cells per cell type does not seem to correlate with the actual distribution of them in the larva. In the case of the 3 dpf dataset, the skeletal putative broad cell type consists of only 66 skeletogenic cells that corresponds to the number of cells present in two larvae at a similar developmental stage (Fig. 3.2). Considering that cells originating from hundreds of larvae and not of a single individual were used for the scRNA-seq experiments, the hypothesis that the computationally produced number of cells does not have any biological meaning is favored. The lower number of skeletal cells represented in the single-cell transcriptome than the number one would expect from a pool of hundreds of larvae could also be linked to the dissociation process, as larval skeletal structures are the last to dissociate, and hence cells can get trapped within the debris and thus be underestimated in our datasets. Nonetheless, the major putative cell types identified as described below are the expected ones, and no known cell type seems to be missing from the analysis, suggesting that the representative cells retrieved are sufficient to address the questions of this study.

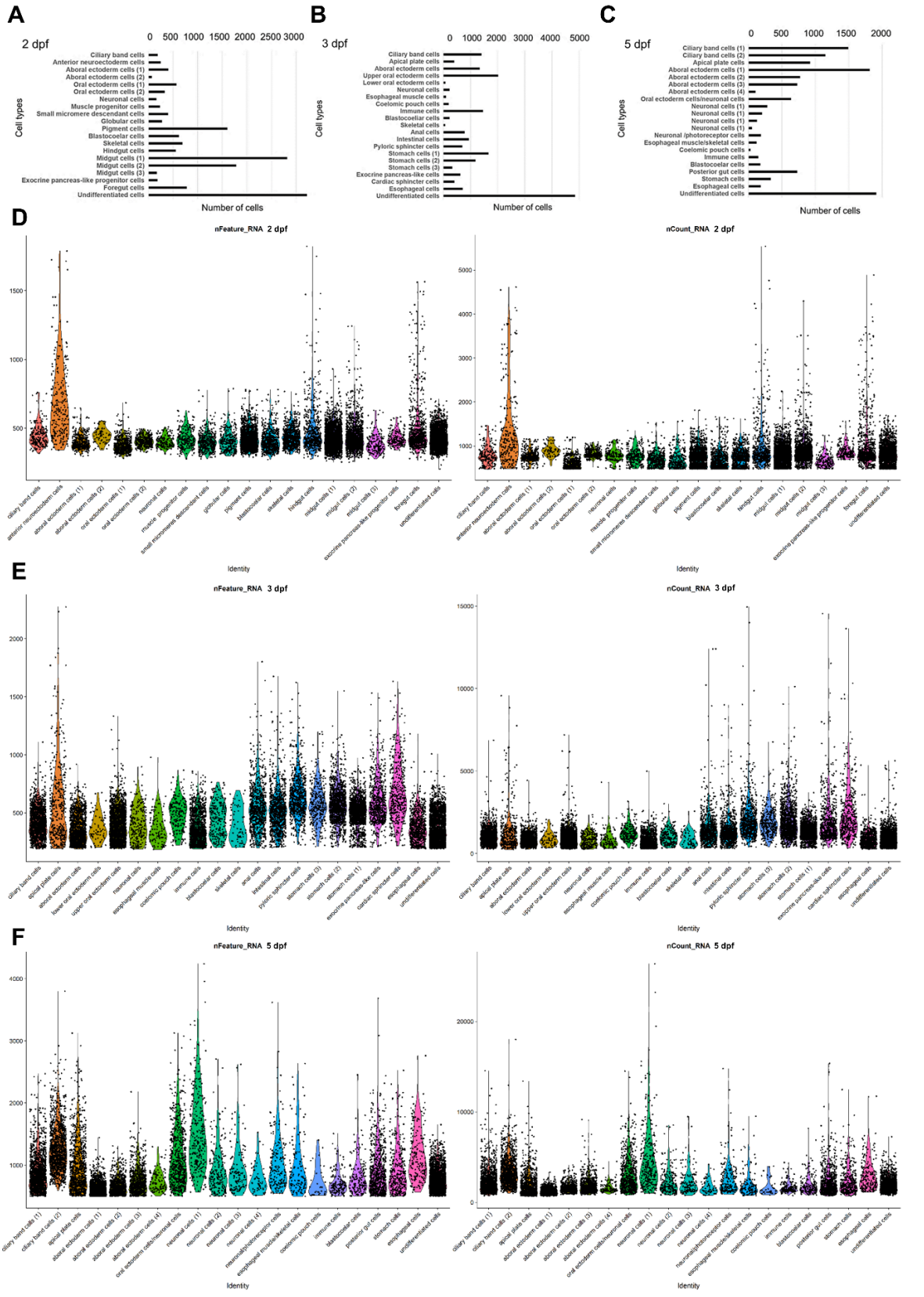


Figure 3.2. Distribution of cells, genes and RNA molecules per putative broad cell type. Bar plots showing the number of cells per cell type at 2 dpf gastrula (A) and 3 (B) and 5 (C) dpf larva stages. Violin plot showing the unique number of genes (nFeature_RNA) and number of molecules (nCount_RNA) at 2 (D) dpf gastrula and 3 (E) and 5 (F) larva stages.

3.1.3. Identification of the putative broad cell types

In order to identify and annotate the computationally produced cell types, a three-step identification process was applied. The first step was to take advantage of the thorough cell type characterization studies performed so far and to use known gene markers as a preliminary way to assess their identity. Next, the identity of cluster markers and of the total genes expressed in every cell type were computationally extracted. Having both the cluster markers, which correspond to the highest and differentially expressed genes, and the total pool of genes expressed, including genes expressed in low levels such as transcription factors, enabled the more confident characterization of the putative broad cell types. This analysis resulted in lists containing 6,365 cluster markers and 15,365 total genes for 2 dpf gastrula, 7,471 cluster markers and 15,578 total genes for 3 dpf pluteus and 5,454 and 16,852 for 5 dpf pluteus stages. These lists can be found in the Non-book component of this thesis (Sp_2dpf_all_genes, Sp_2dpf_marker_genes, Sp_3dpf_all_genes, Sp_3dpf_marker_genes, Sp_5dpf_all_genes, Sp_5dpf_marker_genes). The total number of genes retrieved across all developmental time-points is in agreement with the average of 16,500 genes shown to be expressed by the end of *S. purpuratus* embryogenesis (Tu et al. 2014). The slightly lower number of genes recovered compared to the bulk transcriptomic approaches applied in the past is most probably due to the absence of some genes in the reference transcriptome used or due to minor differences in the developmental timing of the datasets discussed in this thesis compared to the ones produced in Tu et al. 2014. Nonetheless, this slight discrepancy in the number of genes identified is compensated by the cell type specificity our scRNA-seq approach offers.

The last identification step is based on *in vivo* validations of the predicted expression domains of previously characterized and non-characterized genes by performing fluorescent *in situ* hybridizations, immunohistochemical stainings, and combinations of the two. Overall, this process resulted in the identification of 19 out of the 20 gastrula (Fig. 3.4), 20 out of the 21 early pluteus (Fig. 3.5), and 19 out of the 20 (Fig. 3.6) late pluteus stages putative broad cell types.

In all datasets from all three developmental time-points, one cell type could not be traced back to a specific domain of the animal. Due to the high number of cells constituting this cluster compared to the rest (Fig. 3.2), the poorly defined molecular signature as judged by the number of genes found to be expressed (1,563 at 2 dpf,

1,408 at 3 dpf and 1,860 at 5 dpf stages) and lack of specific localization of its gene markers led to its annotation as undifferentiated cells.

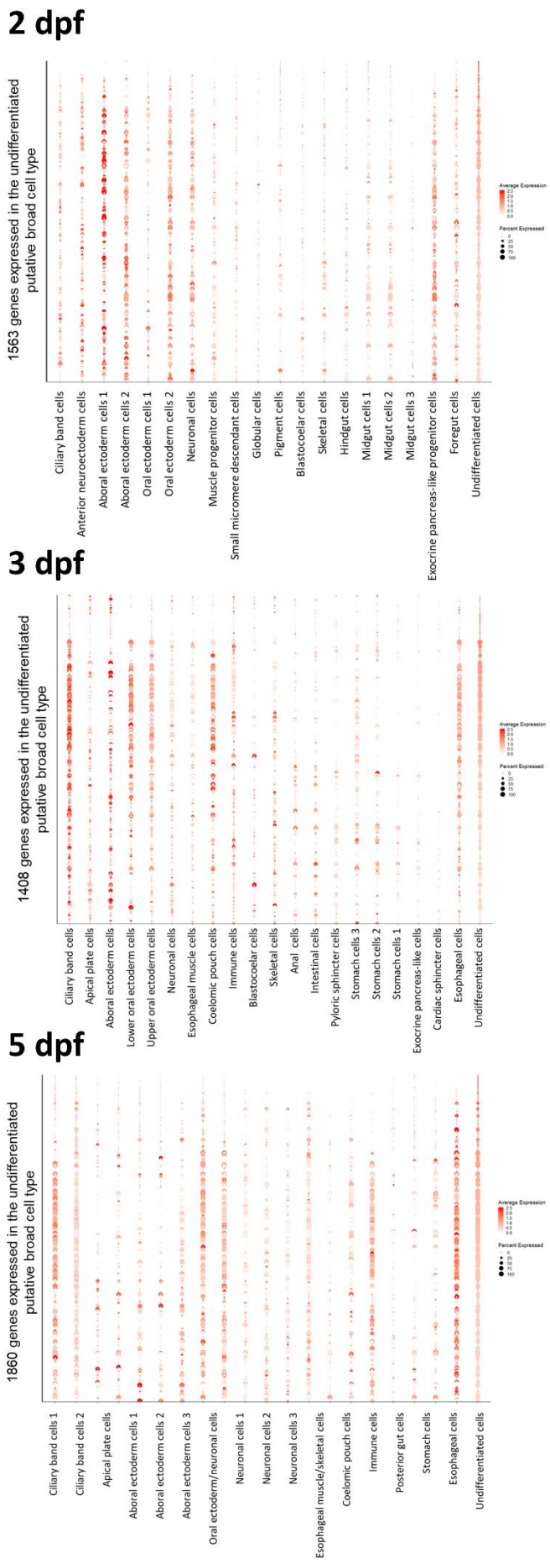


Figure 3.3. Molecular signature of the undifferentiated putative broad cell type. Dotplot of the genes expressed in the undifferentiated putative broad cell type across the three developmental time-points shows high similarity with ectodermally derived cell types. Details on the identity of the genes plotted here can be found in the all_genes files of the non-book component.

Extracting the identity of the total genes expressed in this cell type and in all three developmental time-points resulted in identifying a transcriptomic similarity with ectodermal cell types such as the ciliary band and the oral ectoderm putative broad cell types (Fig. 3.3). Interestingly, a recent study by Perillo et al., in which the authors used scRNA-seq on enriched ectoderm plutei tissues to describe several cell states in place, revealed the presence of a similar uncharacterized ectodermal cell type (Perillo et al. 2020) as the one present in our datasets, thus suggesting that this cell population exists and is not an artifact of the analysis. Moreover, violin plots of the genes (nFeature_RNA) and

molecules (nCount_RNA) per cell type/developmental stage (Fig. 3.2 D, E, F) show that the undifferentiated cell type has a comparable number of molecules to verified cell types. For instance, the undifferentiated cells, ciliary band, immune and esophageal clusters of the 3 dpf, all have similar number of UMIs, supporting the hypothesis that this is a real cell type and not dying cells or empty droplets. Considering the great plasticity and regeneration capability of the sea urchin embryo and larva, it is also possible that this non-differentiated ectodermally associated cell type is a stand by population waiting to be activated.

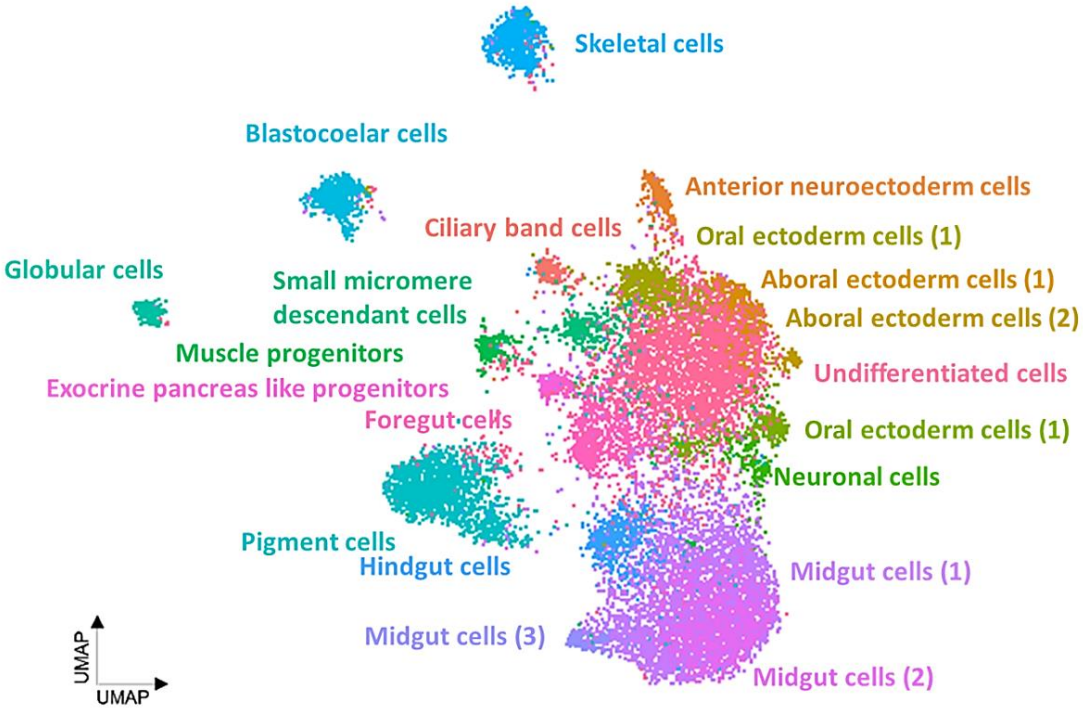


Figure 3.4. UMAP showing the 2 dpf embryonic cell types. Each putative broad cell type is being represented with a different color.

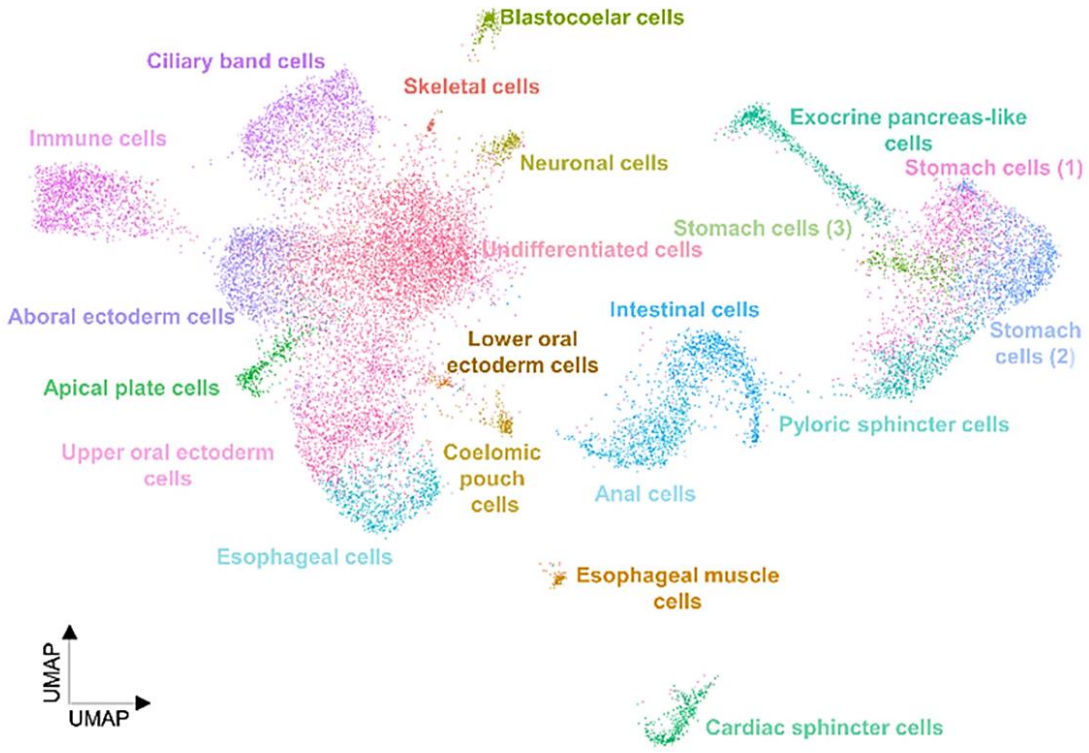


Figure 3.5. UMAP showing the 3 dpf larval cell types. Each putative broad cell type is being represented with a different color.

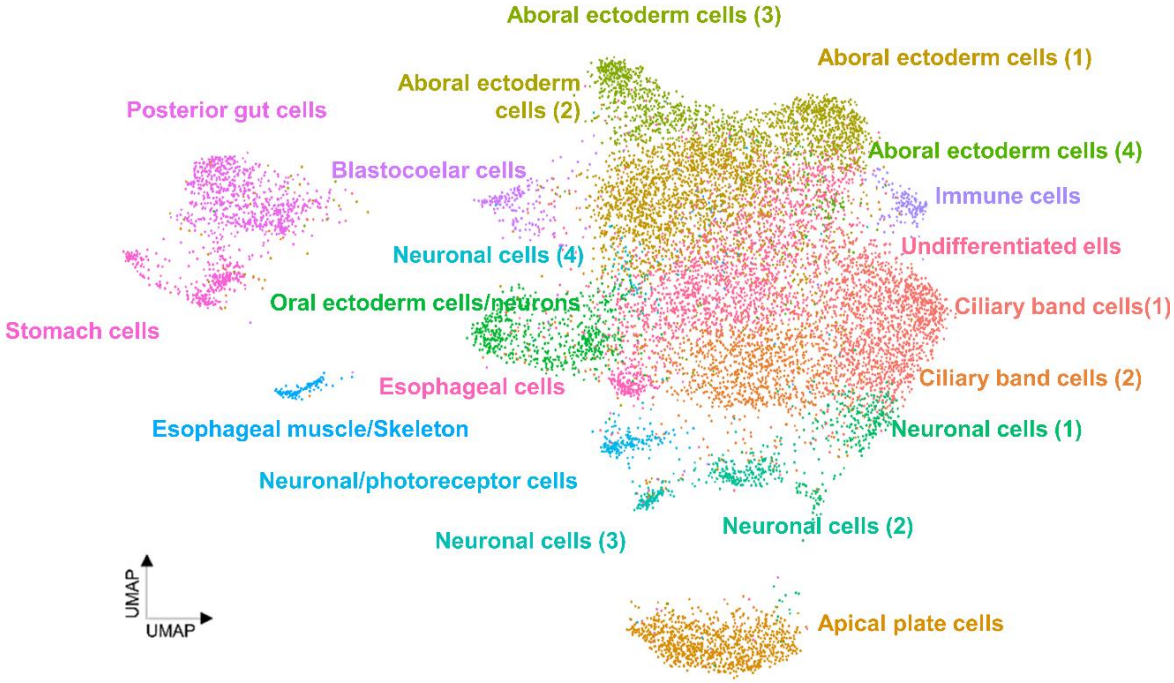


Figure 3.6. UMAP showing the 5 dpf larval cell types. Each putative broad cell type is being represented with a different color.

The use of a mild dissociation protocol, the presence of this type in all replicates and datasets produced by different groups applying different dissociation protocols support the notion that this is a real cell type. Nonetheless, since it does seem to represent an undifferentiated cell type (according to its weak molecular signature), it has been excluded from downstream analysis. Future studies are needed to verify the origin and function of this cell type during sea urchin development.

Plotting the average expression of selected known cell-type gene markers (Fig. 3.5) enabled a preliminary identification of clusters (Figs. 3.7, 3.8, 3.9) that correspond to ciliary band (*Btub2*) (Harlow and Nemer 1987), anterior neuroectoderm/apical plate (*Hbn*) (Burke et al. 2006), aboral ectoderm (*Spec2a*) (Yuh et al. 2001), stomodeum, lower oral ectoderm (*Bra*) (Wei et al. 2012) and upper oral ectoderm (*Gsc*) (Wei et al. 2012), neurons (*SynB*) (Burke et al. 2006), esophageal muscles and muscle-precursors (*Mhc*) (Andrikou et al. 2013), small micromere descendants/coelomic pouches (*Nan2*) (Juliano et al. 2010), blastocoelar cells (185/333) (Ho et al. 2017), immune cells (*Gcm*) (Materna et al. 2013) and skeleton (*Msp130*) (Harkey et al. 1992) across all three developmental time-points.

At 2 dpf gastrula stage, the posterior gut molecular markers *Hox11/13b, Cdx,* and *Pdx1* were predicted to be expressed in the hindgut putative broad cell type, whereas the same markers at 3 dpf pluteus stage were predicted to be expressed in hindgut derived cell types such as anus, intestine, and pyloric sphincter respectively. The same markers were enriched in the 5 dpf posterior gut putative cell type. These predictions recapitulated the experimentally demonstrated expression patterns of all the genes mentioned above. Additional plotting of markers known to label specific domains of the sea urchin embryo and larva gut such as *Chp, ManrC1a, Endo16, Ptf1a, Brn12/4* and *Troponin* (Cole and Arnone 2009, Annunziata and Arnone 2014, Perillo et al. 2016, Yaguchi et al. 2017), resulted in the recognition of three different midgut domains at 2 dpf embryo, three different stomach domains at 3 dpf pluteus, an exocrine pancreas-like domain at gastrula and early pluteus, an esophageal related cell type in all three datasets and one cluster corresponding to cardiac sphincter present only in the 3 dpf pluteus dataset. The same lineage tracing and analysis experiments mentioned previously allowed the grouping of the computationally predicted putative broad cell types according to their developmental origins, enabling a more extensive analysis.

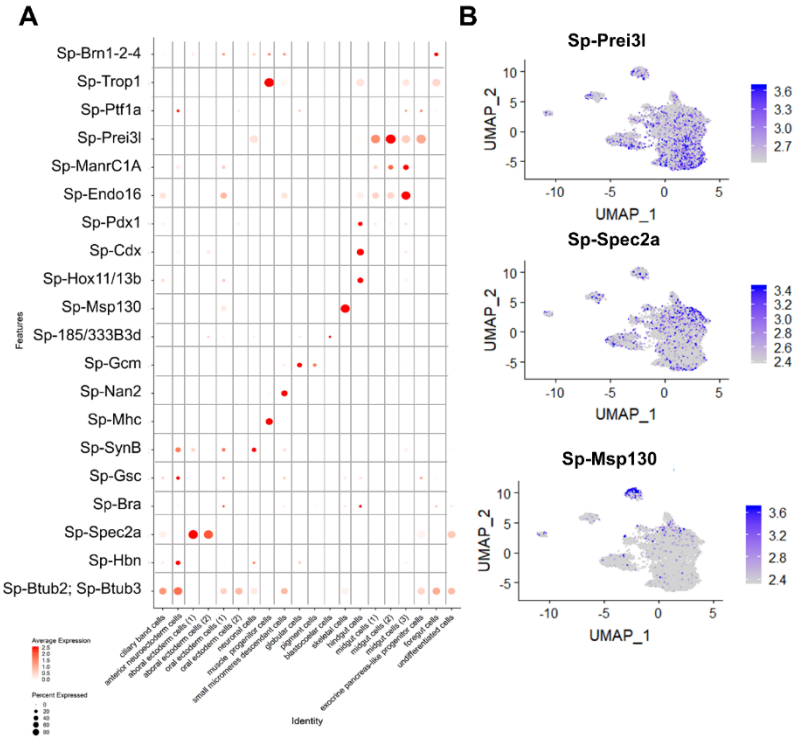


Figure 3.7. Identification of the 2 dpf dataset cell types. A) Dotplot of gene markers previously shown to be cell type-specific. B) Feature plots of specific markers previously shown to be cell type-specific in sea urchin.

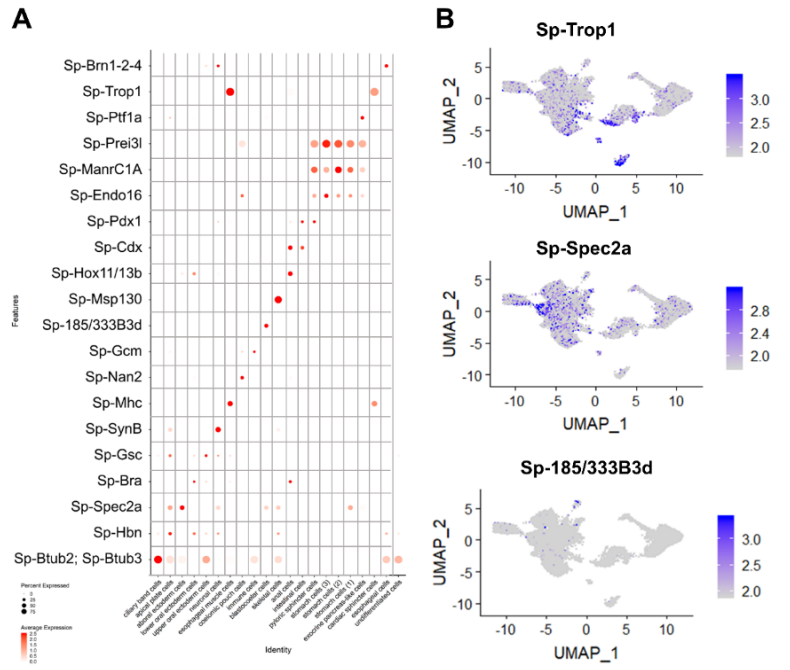


Figure 3.8. Identification of the 3 dpf dataset cell types. A) Dotplot of the same cell type-specific gene markers shown in Figure 3.7. B) Feature plots of specific markers previously shown to be cell type-specific in sea urchin.

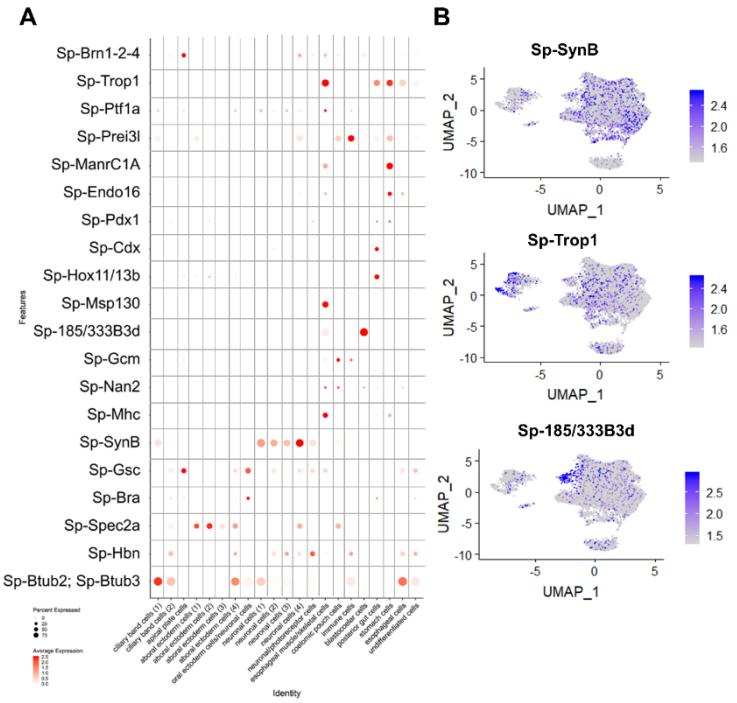


Figure 3.9. Identification of the 5 dpf dataset cell types. A) Dotplot of the same cell type-specific gene markers shown in Figure 3.7. and 3.8. B) Feature plots of specific markers previously shown to be cell type-specific in sea urchin.

Performing FISH and or IHC of characterized and non-characterized marker genes allowed the in *vivo* validation of the scRNA-seq analysis predictions. The high overlap between the computationally predicted gene expression domains and the ones observed *in vivo* (Figs. 3.10, 3.11, 3.12) verified the high quality of the produced datasets and the correct annotation of the cell types and allowed their mapping on distinct embryonic and larval domains (Fig. 3.13, 3.14). For instance, *Sp-Hypp_1249* and *Sp-hypp_2386* are specifically labeling pigment and skeletal cells, at gastrula and early pluteus stages (Figs. 3.10 H, I, 3.11 H, I) respectively, while *Sp-Mlckb* is musculature markers in all time-points tested (Figs. 3.10 J, 3.11 J, 3.12 E).

Concerning the 2 dpf gastrula and 3 dpf early pluteus datasets, all the cell types demonstrated by previous studies to be present are also being recognized by our analysis. Additionally, an increasing cell type complexity as judged by the number and identity of the clusters across these two time-points is identified. This increasing cell type complexity is in line with the morphological differences of these subsequent developmental stages. For instance, one crucial difference is the anatomy of the gut that resembles a straight tube at gastrula stage, whereas at pluteus stage is compartmentalized to three distinct domains as an immediate result of sphincter formation. Furthermore, the larval digestive tract is equipped with cell types controlling its function, such as muscles and neurons. Moreover, the overall shape of the animal changes due to the formation of the ciliary band and skeletal rods.

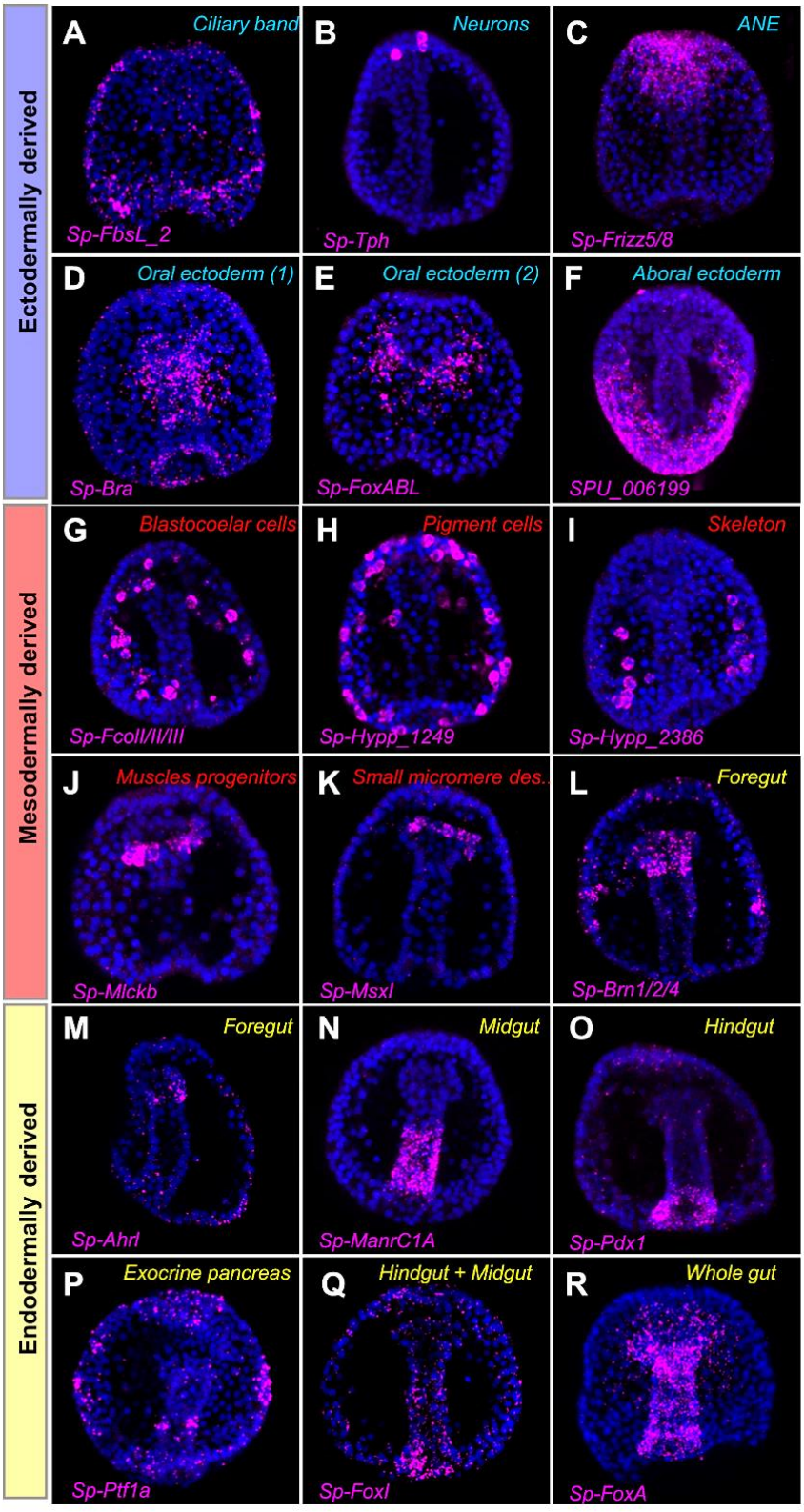


Figure 3.10. FISH validation of the single-cell predicted patterns. FISH of *S. purpuratus* 2 dpf targeting the mRNA for *Sp-FbsL_2* (A), *Sp-Tph* (B), *Sp-Frizz5/8* (C), *Sp-Bra* (D), *Sp-FoxABL* (E), *SPU_006199* (F), *Sp-FcolI/II/III* (G), *Sp-Hypp_1249* (H), *Sp-Hypp_2386* (I), *Sp-Mlckb* (J), *Sp-MsxI* (K), *Sp-Brn1/2/4* (L), *Sp-Ahrl* (M), *Sp-ManrC1A* (N), *Sp-Pdx1*(O), *Sp-Ptf1a* (P), *Sp-Foxl* (Q), *Sp-FoxA* (R)). Color-code indicates the germ layer origins: ectodermally derived, in blue, mesodermally derived, in red, and endodermally derived, in yellow. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections.

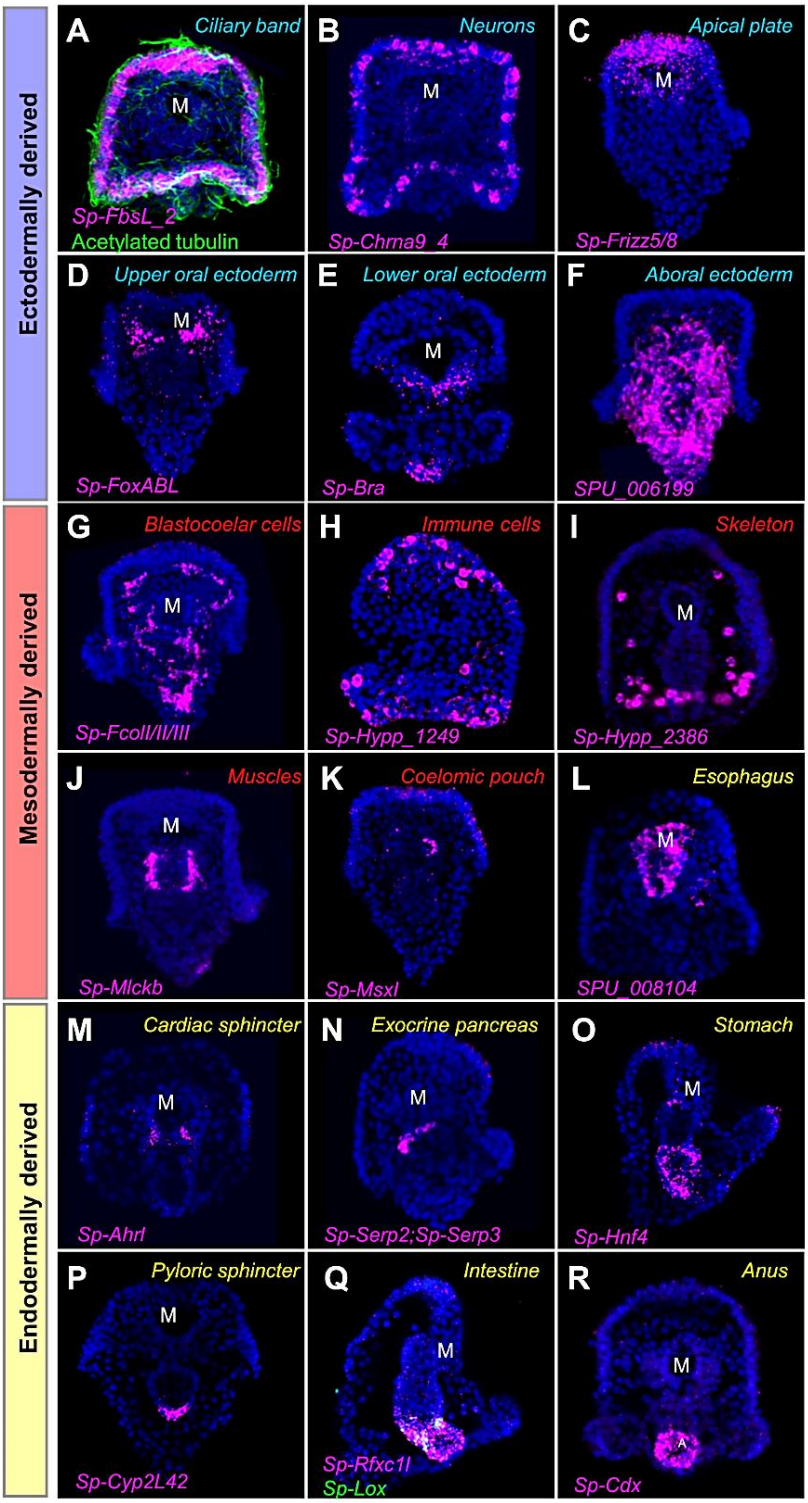


Figure 3.11. Expression pattern of cell type-specific gene markers (3 dpf). FISH with specific antisense probes for the genes Sp-FbsL_2 (A), Sp-Chrna9_4 (B), Sp-Frizz5/8 (C), Sp-FoxABL (D), Sp-Bra (E), SPU_006199 (F), Sp-FcolI/II/III (G), Sp-Hypp_1249 (H), Sp-Hypp_2386 (I), Sp-Mlckb (J), Sp-Msxl (K), SPU_008104 (L), Sp-Ahrl (M), Sp-Serp2; Sp-Serp3 (N), Sp-Hnf4 (O), Sp-Cyp2L42 (P), Sp-Rfxc1l (Q), Sp-Pdx1 (Q) and Sp-Cdx (R).

Immunofluorescent detection of acetylated tubulin (in green) is used to highlight that the Fbsl_2 positive domain (A) is the ciliary band. Color-code indicates germ layer origins: ectodermally derived, in blue, mesodermally derived, in red, and endodermally derived, in yellow. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections.

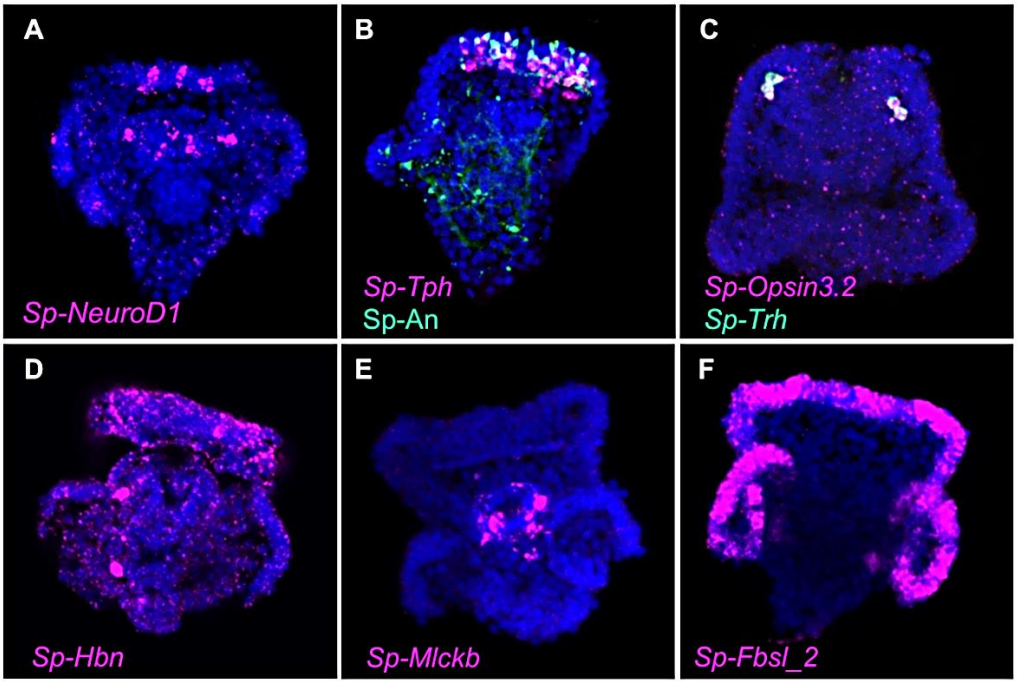


Figure 3.12. Validation of single-cell predictions for the 5 dpf dataset. FISH with specific antisense probes for the genes *Sp-NeuroD1* (A), *Sp-Tph* (B), *Sp-Opsin3.2* (C), *Sp-Trh* (D), *Sp-Hbn* (D), *Sp-Mlckb* (E), *Sp-Fbsl_2* (F), *Sp-Hypp_1249* (H), *Sp-Hypp_2386* (I), *Sp-Mlckb* (J), *Sp-Msxl* (K), *SPU_008104* (L), *Sp-Ahrl* (M), *Sp-Serp2; Sp-Serp3* (N), *Sp-Hnf4* (O), *Sp-Cyp2L42* (P), *Sp-Rfxc1l* (Q), *Sp-Pdx1* (Q) and *Sp-Cdx* (R). Immunofluorescent detection of the neuropeptide Sp-An (B). Color-code indicates germ layer origins: ectodermally derived, in blue, mesodermally derived, in red, and endodermally derived, in yellow. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections.

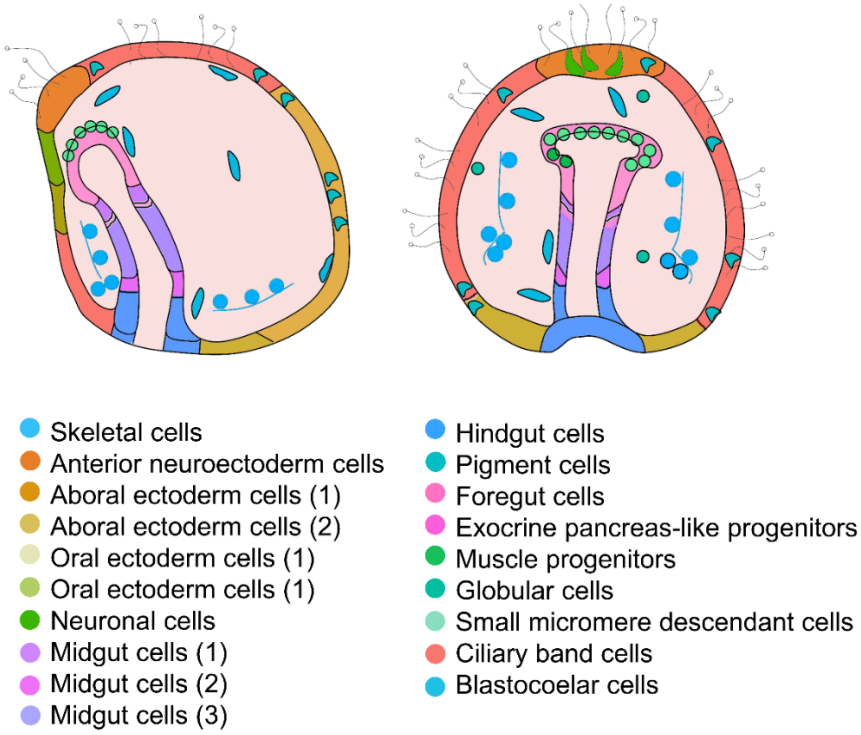


Figure 3.13. Cartoon showing the location of the identified cell types on different embryonic domains. Color-code is the same as in Figure 3.4.

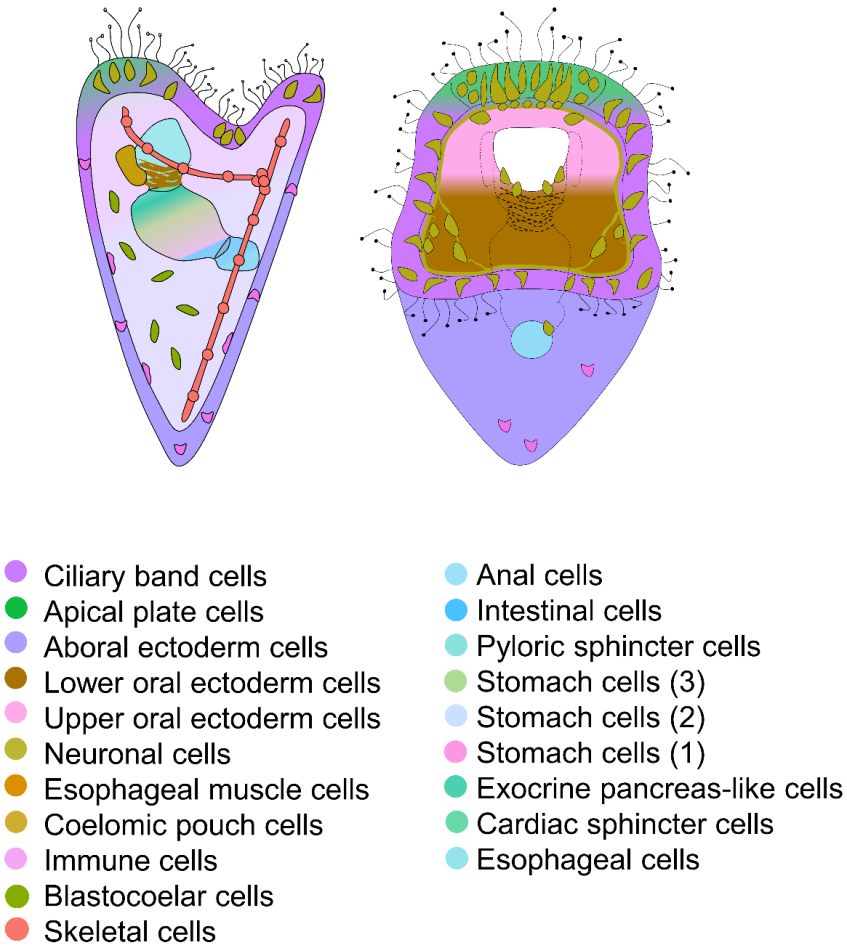


Figure 3.14. Cartoon summarizing the location of the identified cell types on different larval domains (3 dpf). Color-code is the same as in Figure 3.5.

Interestingly, these changes can be traced in the scRNA-seq data. In detail, the gastrula scRNA-seq atlas contains 6 endodermally derived putative broad cell types, while the 3 dpf pluteus stage atlas consists of 9 cell types deriving from the endoderm. Two out of the three additional computationally generated cell types correspond to the pyloric and cardiac sphincters, a larval stage innovation. On the other hand, the third additional cell type derives from the differentiation of hindgut putative broad cell type into the intestine and anus' specialized domains.

One interesting observation was that although gastrula and early pluteus stages are morphologically very different, consisting of different putative broad cell types, when the 2 and 3 dpf datasets were computationally integrated, similar cell types could be identified as if the 3 dpf dataset was analyzed alone (Figs. 3.15, 3.16). The only reported difference is that the cross-stages integration resulted in two additional immune system cell types compared to the 3 dpf larva's cell types only. In detail, the 3 dpf dataset contains a single immune system putative broad cell type, while the 2

dpf one contains two corresponding to globular and pigmented cells. To understand whether these additional putative broad cell types in the cross-stages integrated dataset reflect the immune cell types found at gastrula stage or are present in both developmental points and their appearance as distinct cell types is a result of the increase of the initial amount of cells analyzed, further analysis was performed (Fig. 3.17). Subclustering analysis of the single immune system cell type present at 3 dpf pluteus and plotting of known gene markers labeling distinct sea urchin immune system types resulted in the identification of 8 distinct immune system populations (Fig. 3.17 A, B), with two of the corresponding to pigment cells and globular cells as judged by the expression of the markers *Sp-Pks1* and *Sp-MacpfA2* respectively (Ho et al. 2017). To confirm whether the subclustering of a putative broad cell type is sufficient to predict subtypes accurately, a similar analysis was also performed on one of the best-characterized sea urchin cell types, the early pluteus skeletal cell type. It has been previously shown that PMCs along the syncytium have non-universal patterns of genetic programs, and based on this, they were grouped into five classes (Sun and Ettensohn, 2014). Interestingly, the subclustering analysis of the 3 dpf pluteus skeletal broad cell type revealed five PMCs populations with distinct molecular fingerprint (Fig. 3.18), in line with that previous observation, building confidence that subclustering can reveal diverse subtypes.

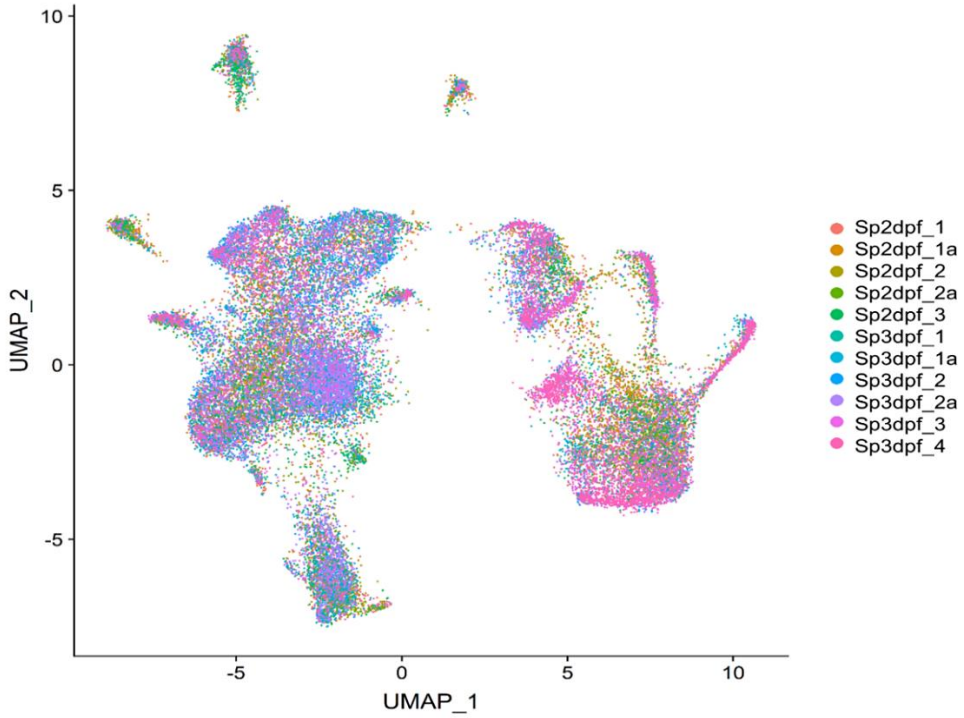


Figure 3.15. Data integration across developmental stages. UMAPs overlay by original identity showing the overlap of the different biological and technical replicates for 2 dpf and 3 dpf developmental stages.

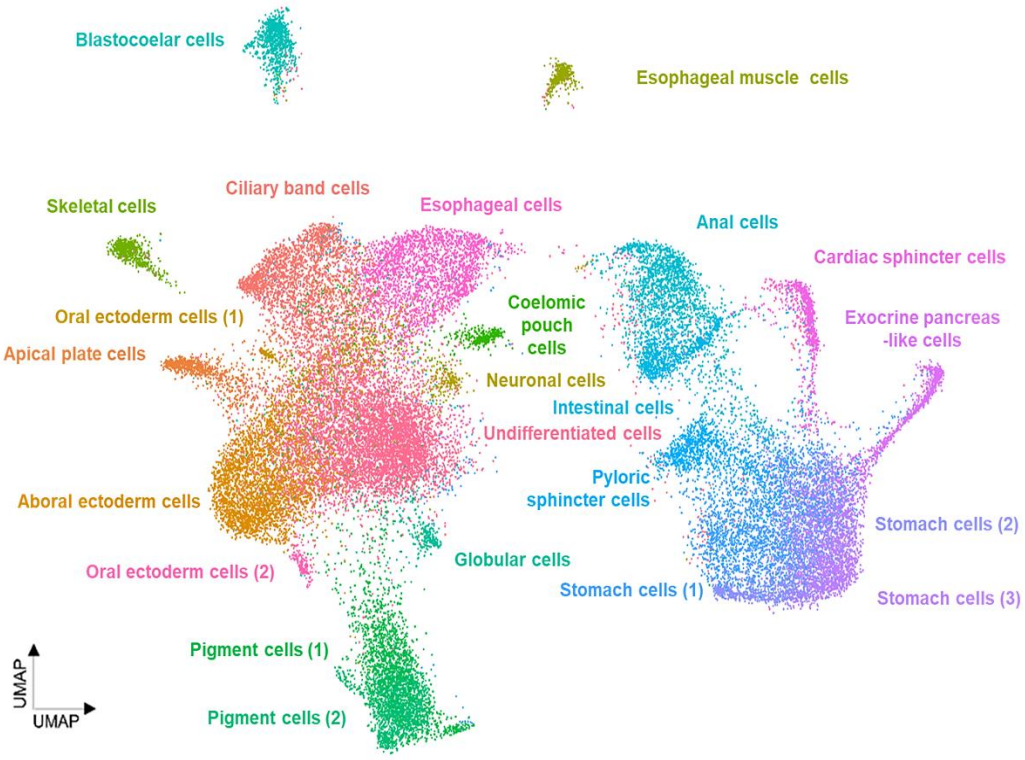


Figure 3.16. UMAP showing the common between gastrula and pluteus putative broad cell types. Each putative broad cell type is being represented with a different color.

Selecting the 3 immune system clusters present in the merged dataset and plotting the objects against their original identity showed that cells from both developmental time-points participate in forming those clusters (Fig.3.17 C). This observation leads to the conclusion that the increased amount of cells used for the analysis when merging the datasets from the two different time-points influences the clustering allowing the identification of transcriptomic differences that otherwise could only be revealed by subclustering. It also confirms the choice of using the term putative broad cell type since this clearly demonstrates that computationally produced clusters do not correspond to 1:1 cell types *in vivo*. Nonetheless, the observation that cells from two morphologically different developmental stages share enough transcriptomic similarities allowing them to cluster, overlap, and form similar clusters, suggests that gastrula cells acquire the transcriptomic signature present in plutei cell types before any morphological difference is evident.

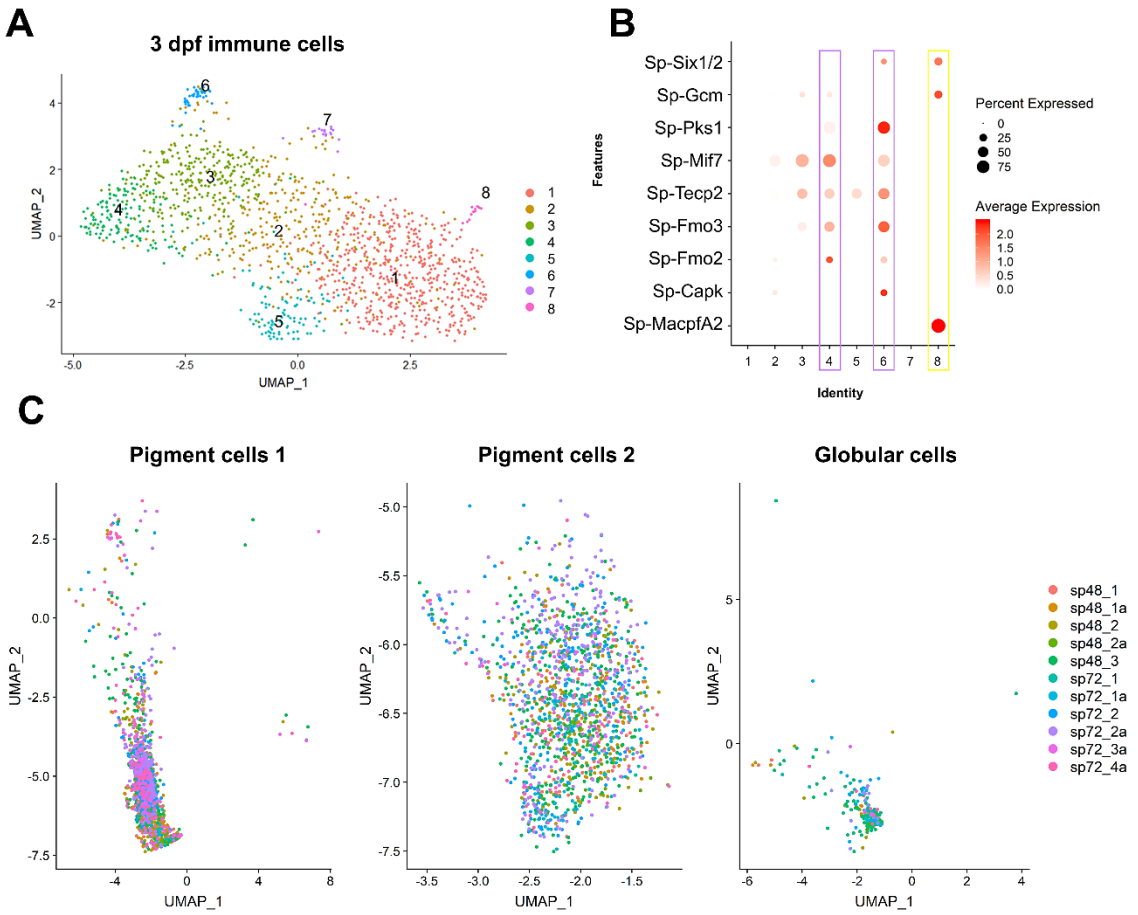


Figure 3.17. Comparison of the immune cell population. A) UMAP showing 8 different immune cells populations at 3 dpf larval stage. B) Dotplot showing gene markers known to be differentially expressed in distinct immune cell populations. Pigment cells (purple) and globular (yellow) cells gene markers are highlighted. C) UMAPs overlay by original identity showing the cellular composition of the Pigment cells 1, Pigment cells 2, and Globular cells that were generated by integrating the 2 and 3 dpf datasets.

Another interesting observation concerns the cell type composition of the 5 dpf pluteus dataset. Surprisingly, although the main cell types expected to be present such as immune cells, skeleton, neurons, muscles, and epithelial cells (ciliary band, apical plate, oral and aboral ectoderm) are present, major cell types patterning the digestive tube such as specialized stomach domains and sphincter cells are missing. Considering that these specialized cell types are present in the 3 dpf dataset and that are essential components of the digestive tract, their absence from the 5 dpf dataset is most likely due to technical issues and not related to the 5 dpf pluteus's biology. An explanation for this could be that for all scRNA-seq libraries constructed; fasted larvae were used to ensure no algal contamination from the food would be present. As mentioned in chapter 1, the feeding process is initiated once the larva's mouth is formed, approximately between 66-72 hours post-fertilization (hpf). As a result, the absence of food can only have a biological effect only after

3dpf. Taken together, this suggests that the lack of differentiated digestive tube cell types in the 5dpf pluteus dataset might be a direct result of the larvae being fastened and that gut cell types began to deteriorate.

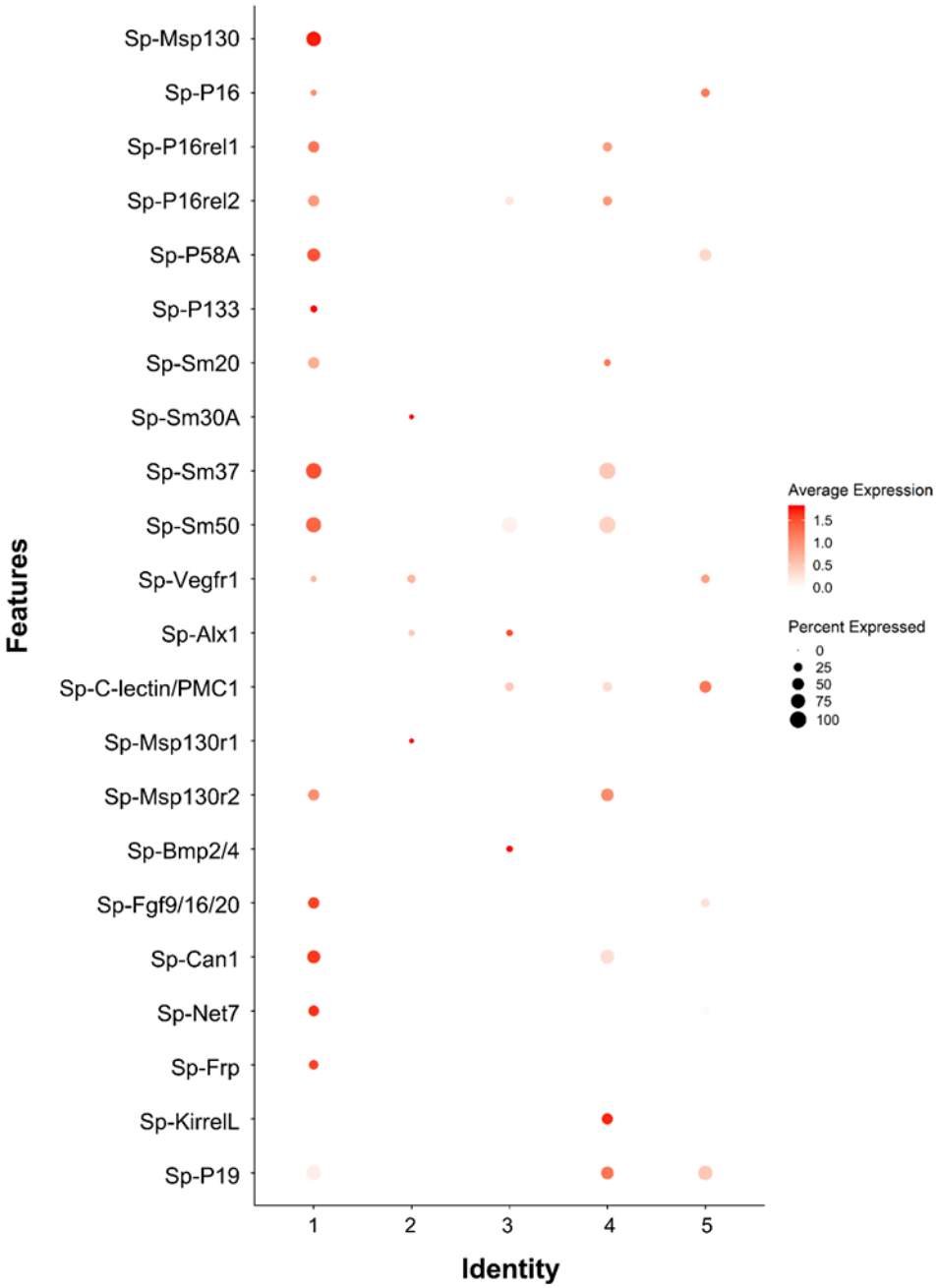


Figure 3.18. Skeletal subclustering at 3 dpf pluteus stage. Dotplot showing the average expression of 22 skeletal gene markers and their distribution among 5 skeletal subtypes, as demonstrated in Sun and Ettensohn, 2014.

Another simple explanation may be that capturing and sequencing more biological replicates is needed to fully reconstruct the cell types in place as cells from only two biological replicates were used. The total number of the 5 dpf pluteus is estimated between 3,000 to 4,000 cells, which is a much greater number compared to the

1,500 cells estimated to be present in the 3 dpf early pluteus, suggesting that more sampling is needed to extract the same level of information as for the rest of the developmental stages. Due to this uncertainty on the exact cell type composition, 5 dpf dataset was excluded from further analysis.

3.1.4. ScRNA-seq as a tool to identify novel expression patterns

ScRNA-seq data can be used as a tool to identify complex and rare gene expression patterns across cell types. In order to assess the power of the scRNA-seq approach, two different strategies were followed. The first one was to choose genes with well-known expression patterns and examine which are the scRNA-seq predicted domains. The second strategy was based on usage of genes described to be functioning or expected to be present in a given cell type, but with low expression levels and thus undetectable by *in situ* hybridization.

The already described neuronal gene markers *Sp-Sip1*, *Sp-SoxC*, *Sp-Hbn,* and the skeletal gene marker *Sp-Fgf9/16/20* are known to be present at 3 dpf larva were chosen. Our scRNA-seq analysis of this developmental stage, combined with the high-resolution FISH protocol developed in the lab in which this thesis was carried out, allowed the identification of novel expression domains for those genes. These novel expression patterns were either never described before or were considered to be a non-specific signal.

The scRNA-seq analysis predicts the transcription factors *Sp-SoxC* and *Sp-Hbn*, both shown to be involved at the early steps of neuronal specification (Garner et al., 2016, Wei et al., 2016, Yaguchi et al., 2016), to be present in neurogenic territories as well as in novel domains such as skeletal cells (Fig. 3.19 A). Indeed, *in situ* hybridization of *Sp-SoxC* and *Sp-Hbn* verified their expression in PMCs (Fig.3.19 B). Furthermore, FISH for *Sp-Hbn* combined with immunostaining using the anti-msp130 antibody, which labels PMCs explicitly along the skeletal rods (Harkey et al., 1992), verified the expression of *Sp-Hbn* in skeletal cells (Fig. 3.19 C). *Sp-Sip1* has been used as a positive control since it has been previously demonstrated in another sea urchin species to have an intricate expression pattern spanning from neurons to skeleton (McClay et al., 2018). Likewise, the FGF signaling ligand, *Sp-Fgf9/16/20*, is involved in the skeletal formation and expressed in specific PMCs populations (Adomako-Ankomah and Ettensohn, 2014). Similarly, scRNA-seq

verified the expected skeletal expression domain while added oral ectoderm, cardiac sphincter, intestine, and anus as additional expression domains (Fig. 3.19). Therefore, it is evident that scRNA-seq can predict complex gene expression patterns previously challenging to detect by *in situ* hybridization.

Based on the observation that scRNA-seq can predict complex and novel expression patterns, I set out to investigate the detection limits by *in situ* hybridization versus scRNA-seq. To this end, two case studies were chosen. The first concerned the maternal transcription factor *Meis*, which was shown to regulate the patterning along the embryo's anterior/posterior axis and to be essential for maintaining the neurogenic ectoderm fate in the sea urchin species *Hemicentrotus pulcherrimus* (Yaguchi et al. 2018). Although an exact role in regulating neurogenesis and determining the anterior neuroectodermal fate was demonstrated, its low expression levels did not allow the localization of *Meis* transcripts by *in situ* hybridization in any specific domain. Plotting for the average expression of *Sp-Meis* across the three developmental points analyzed in this thesis led to identifying specific cell types expressing it. In detail, *Sp-Meis* transcripts were detected in various cell types across development, including the expected but not demonstrated so far anterior neuroectoderm and apical plate putative broad cell types (Fig. 3.20).

The second case study concerned the 3 dpf coelomic pouch cell cluster. The coelomic pouch is the mesodermal part of the larva that will grow into the rudiment and give rise to the juvenile sea urchin after metamorphosis (Strathmann 1987, Smith et al. 2008). The formation of the coelomic pouch is a complex and highly regulated procedure, and it has been demonstrated that small micromeres contribute to this process and the juvenile itself (Strathmann, 1987). In an attempt to characterize this population, a study by Juliano et al. demonstrated that several genes involved in germline determination and maintenance in a variety of organisms are present in the small micromeres and the coelomic pouch of the sea urchin embryo (Juliano et al., 2006). On the other hand, other germ line gene candidates tested by *in situ* hybridizations were either not expressed or enriched in this cell type.

Interestingly, when the average expression of candidate genes from those studies (Juliano et al. 2006, Luo and Su 2012, Martik and McClay 2015) were plotted, their transcripts were found enriched in the coelomic pouch cell cluster (Fig. 3.21).

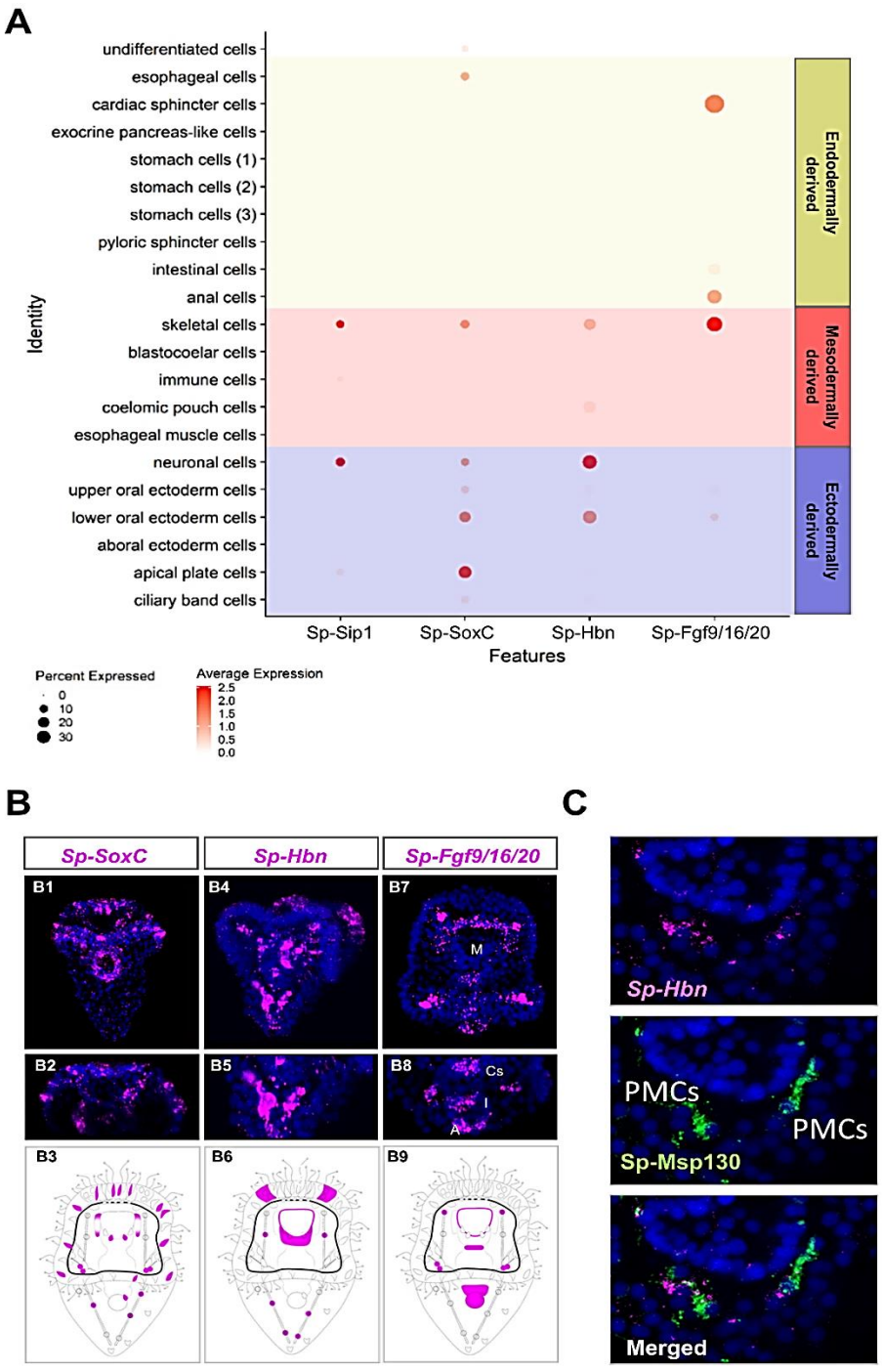


Figure 3.19. Novel expression patterns of known gene markers at 3 dpf pluteus stage. A) Dotplot showing the average expression of *Sp-Sip1*, *Sp-SoxC*, *Sp-Hbn,* and *Sp-Fgf9/16/20* validating the already described expression patterns while predicting novel ones. The putative broad cell types are grouped according to their developmental origins. B) FISH for the genes *Sp-SoxC* (B1-B2), *Sp-Hbn* (B4-B5), and *Sp-Fgf9/16/20* (B7-B8) in line with the scRNA-seq predictions. Cartoons depicting the overall expression domains of *Sp-SoxC* (B3), *Sp-Hbn* (B6), and *Sp-Fgf9/16/20* (B9). C) FISH using antisense probe for Sp-*Hbn* paired with immunostaining for Msp130. DAPI labels the nuclei (in blue). All images are stacks of merged confocal Z sections. A, Anus; Cs, Cardiac sphincter; M, Mouth; I, Intestine; PMCs: Primary mesenchyme cells; Ps, Pyloric sphincter; St, Stomach.

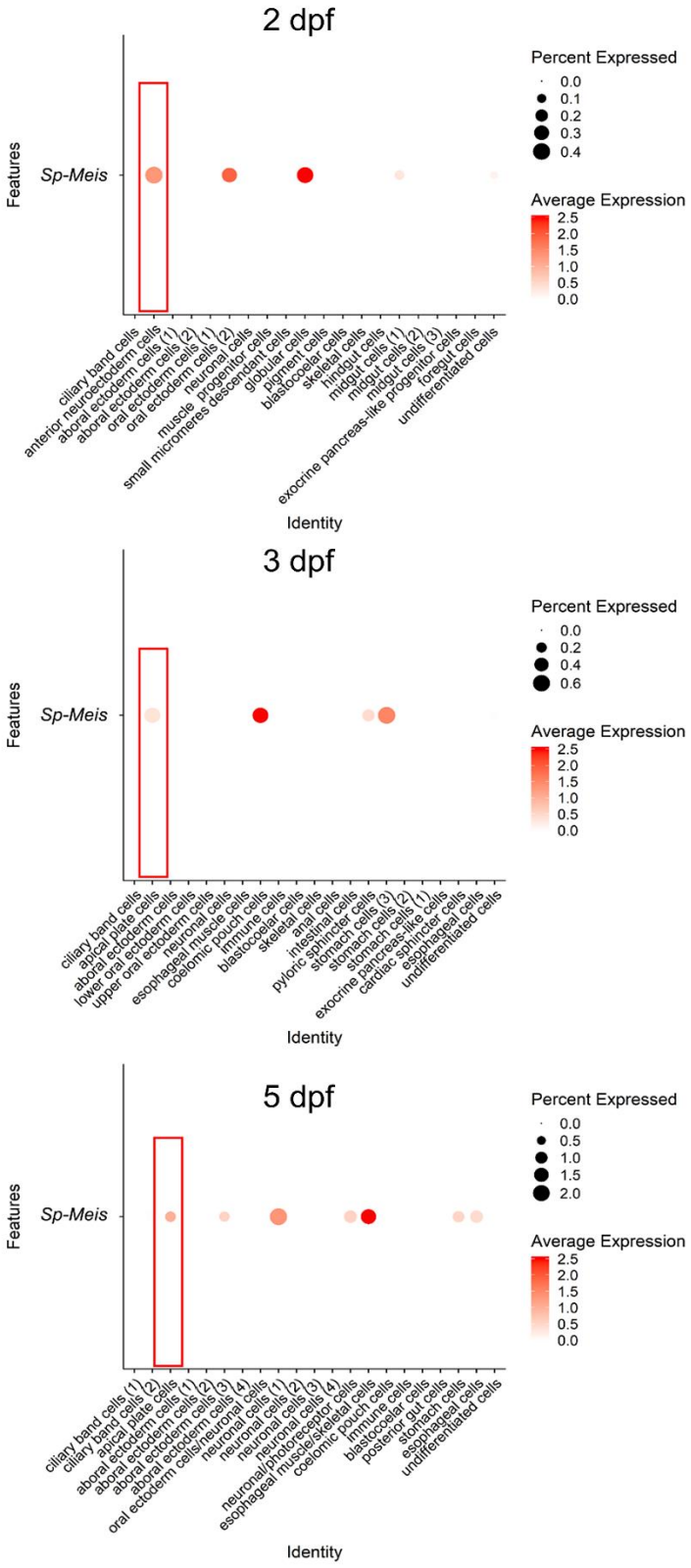


Figure 3.20. ScRNA-seq reveals the expression of a gene previously undetectable by *in situ* hybridization. Dotplot is showing the expression domain of the transcription factor *Meis.*

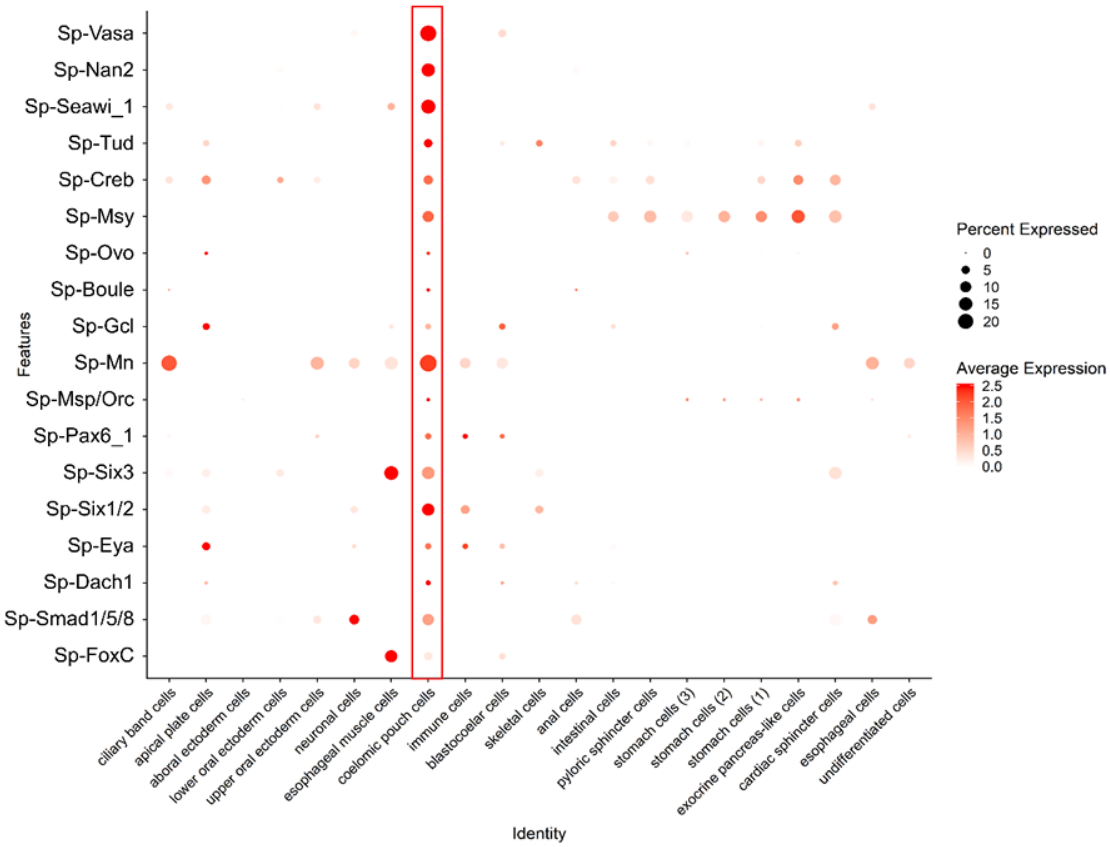


Figure 3.21. ScRNA-seq reveals novel expression patterns of genes. Dotplot of genes known for their involvement in germline determination in other species are all enriched in coelomic pouch cells.

Taken together, the analyses concerning the expression pattern prediction for known genes, and previously undetectable by *in situ* hybridization genes suggest that i) scRNA-seq data can accurately predict gene expression patterns, ii) combination of scRNA-seq and visualization techniques such as immunohistochemistry and FISH can be used to reveal novel gene expression patterns, and iii) that scRNA-seq technology offers higher detection ability compared to the *in situ* hybridization.

3.1.5. Characterization of the early pluteus larva

Since this initial analysis allowed the successful recognition of the cell types present in all three datasets, the focus was shifted towards further analyzing the 3 dpf early pluteus dataset. The reasons for this lie in the lack of extensive regulatory and molecular analysis available at this developmental time-point and the presence of fully differentiated cell types. Moreover, this is a milestone stage of sea urchin's life cycle by being the first free-swimming and feeding larval stage.

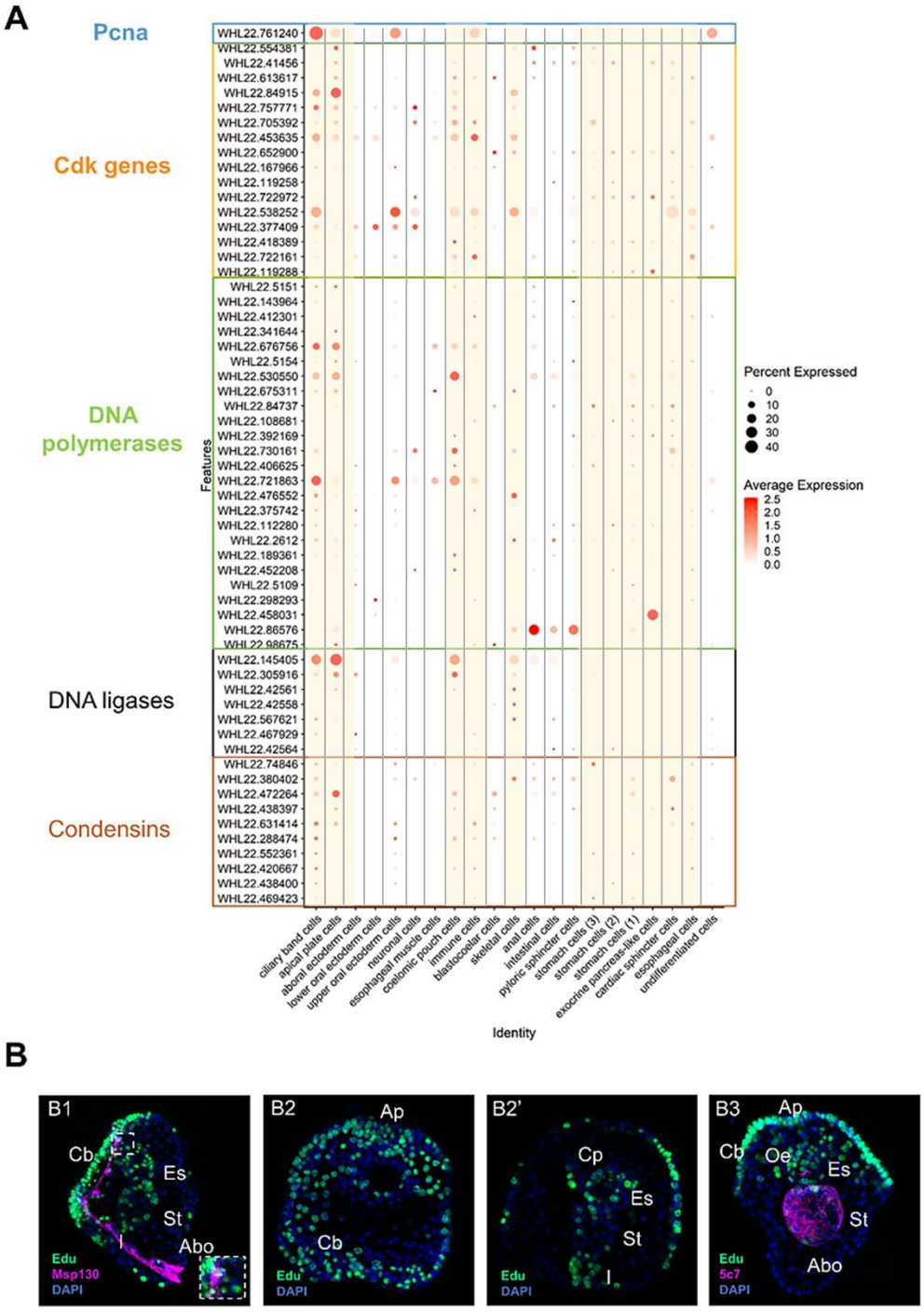


Figure 3.22. Proliferation status and dynamics of the larval cell types. A) Dotplot showing the average expression of genes encoding for *cdk1, pcna*, DNA polymerases, DNA ligases, condensins, and centromere proteins. Yellow boxes indicate highly proliferative cell types. B) Proliferating cells, as indicated by EdU labeling. B1) 3 dpf larva stained with EdU and the skeletogenic marker Msp130. B2) Selected confocal sections of a larva in oral view labeled with EdU. B2') Different confocal sections showing the same larva as in B2 in oral view stained with EdU. B3) Larva labeled with EdU and the endodermal marker 5c7. Abo: Aboral ectoderm; Ap: Apical plate; Cb: Ciliary band; Es: Esophagus; Oe, Oral ectoderm; St: Stomach

Most differentiated cell types can proliferate to replenish cells that have been lost due to injury or cell death, while some lose the proliferation ability once differentiated. To investigate which putative broad sea urchin cell types are actively proliferating, I plotted genes such as *pcna,* DNA polymerases, DNA ligases, condensins, and centromere proteins shown to be involved in sea urchin cell proliferation (Perillo et al., 2020). The majority of those genes were enriched in the ciliary band, apical plate, coelomic pouch, immune, skeletal cell clusters, and several endodermally derived cell types (Fig. 3.22 A). Validation of those predictions was performed using Edu pulse labeling for 2 hours, followed by immunohistochemical detection of the endodermal marker 5c7 (Wessel and McClay 1987) and the skeletogenic cells marker Msp130 (Fig. 3.22 B). That set of experiments revealed that S-phase cells were detected in the same cell types, as predicted by the scRNA-seq data, while cell types predicted not to proliferate, such as the aboral ectoderm cell cluster, did not show any Edu related fluorescence.

3.2. Identification of different regulatory states at 3 dpf

The transcription factor content is what provides and maintains the identity of a given cell type. To unravel the regulatory states of the identified putative broad cell types, a comparison of the different transcription factor cocktails present within the different cell types was performed. Doing so, all the different broad cell types were merged to artificially generate the three germ layers from which it has been previously demonstrated they come from (Cameron et al., 1987, McClay, 2011).

Due to the stereotypic development of the sea urchin embryo/larva and taking into account a recent study performed by Sladitschek and co-authors, demonstrating that the single cell transcriptomes contain information about the lineage from which individual cells come from (Sladitschek et al., 2020), it was hypothesized that the same applies in this case. Moreover, a ClusterTree analysis, incorporated in the Seurat clustering analysis computational package, was performed in order to compare the different cell types (Fig. 3.23). ClusterTree compares the different cell types and shows this comparison's results in the form of a tree. The resulting tree reflects the relationship of an 'average' cell from each identity class and is estimated based on a distance matrix constructed in gene expression space. The tree generated here suggested that most of the endodermally derived cell types seem to

be highly related, while anal and intestinal clusters seem to cluster more with ectodermal and mesodermal putative broad cell types. Nonetheless this tree shows that several developmentally linked cell types such as stomach, pyloric, exocrine and cardiac sphincter coming from the endoderm as well as apical plate, aboral ectoderm, ciliary band and upper oral ectoderm cells coming from the ectoderm seem to be transcriptionally similar to each other.

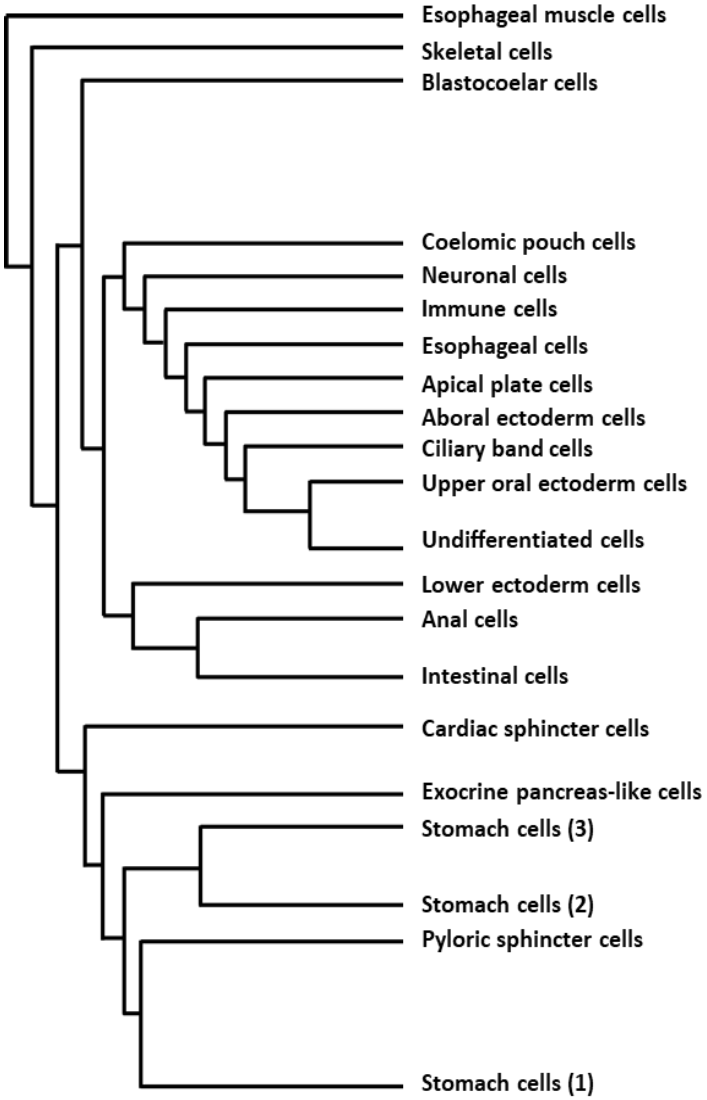


Figure 3.23. Several cell type developmental origins can be reflected by scRNA-seq, ClusterTree showing the grouping of the putative cell types.

Comparing the transcription factor content of the cell types deriving from ectoderm, mesoderm and endoderm revealed that only a few of them are germ layer-specific, and the majority of them seems to be expressed in derivatives of more than one germ layer. The three germ layers share 187 TFs, approximately 1/3 of their transcription factor toolkit (Fig. 3.24). Interestingly, mesodermal cell types seem to share a higher number of transcription factors with ectodermal rather than endodermal ones, with which they have tightly linked developmental origins. Depending on the comparison, cell types derived from the same germ layer seem to share a significant amount of TFs (1/3 in most cases), whereas most of them seem to be cell type-specific (Fig. 3.24 B, D & E).

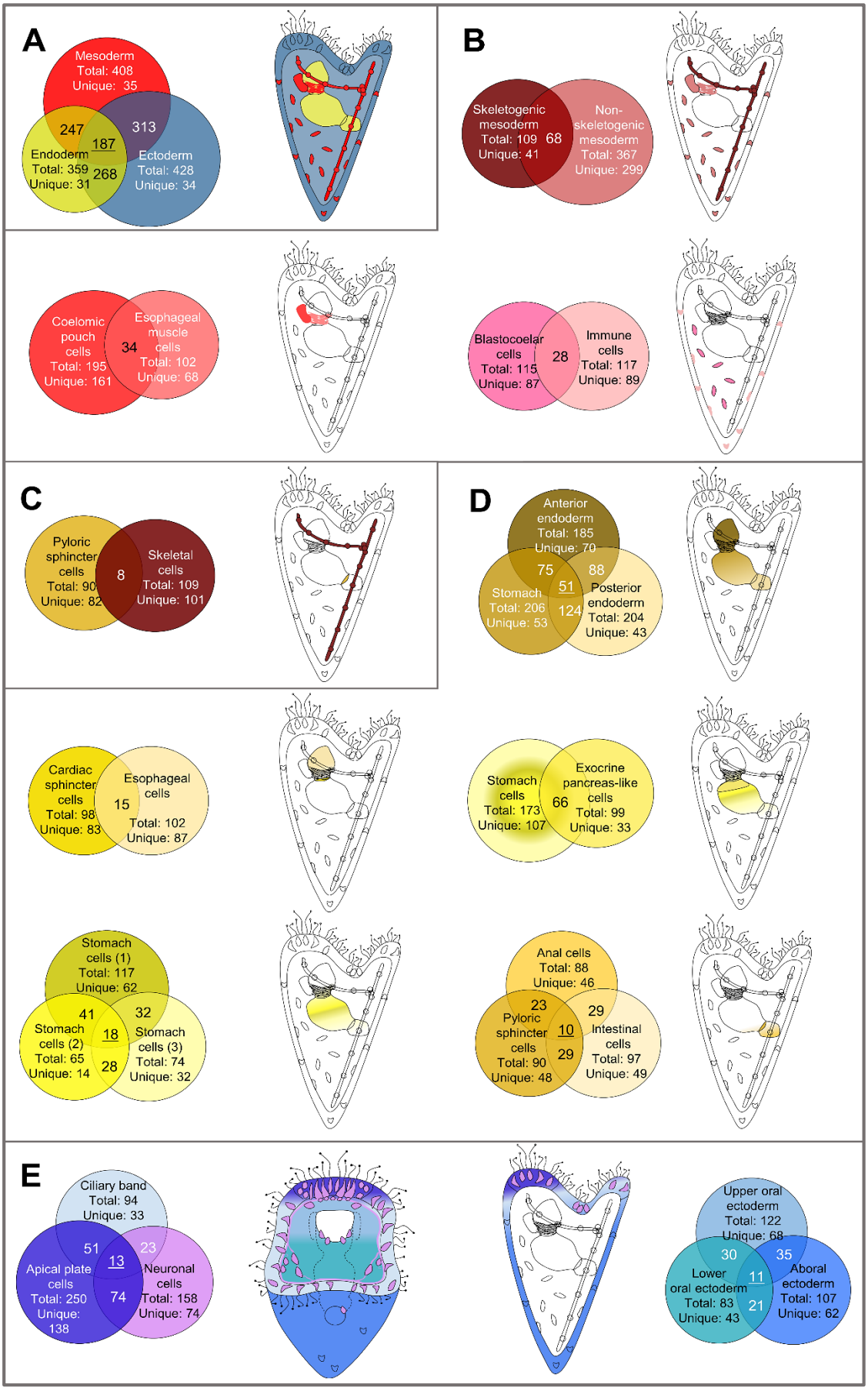

A
Mesoderm
Total: 408
Unique: 35
247
313
187
Endoderm
Total: 359
Unique: 31
268
Ectoderm
Total: 428
Unique: 34
B
Skeletogenic
mesoderm
Total: 109
Unique: 41
68
Non-
skeletogenic
mesoderm
Total: 367
Unique: 299
Coelomic
pouch
cells
Total: 195
Unique: 161
Esophageal
muscle
cells
34
Total: 102
Unique: 68
Blastocoelar
cells
Total: 115
Unique: 87
28
Immune
cells
Total: 117
Unique: 89
C
Pyloric
sphincter
cells
Total: 90
Unique: 82
8
Skeletal
cells
Total: 109
Unique: 101
D
Anterior
endoderm
Total: 185
Unique: 70
75
51
88
Stomach
Total: 206
Unique: 53
124
Posterior
endoderm
Total: 204
Unique: 43
Cardiac
sphincter
cells
Total: 98
Unique: 83
Esophageal
cells
15
Total: 102
Unique: 87
Stomach
cells
Total: 173
Unique: 107
66
Exocrine
pancreas-like
cells
Total: 99
Unique: 33
Stomach
cells (1)
Total: 117
Unique: 62
41
32
Stomach
cells (2)
Total: 65
Unique: 14
18
28
Stomach
cells (3)
Total: 74
Unique: 32
Anal cells
Total: 88
Unique: 46
23
29
Pyloric
sphincter
cells
Total: 90
Unique: 48
10
29
Intestinal
cells
Total: 97
Unique: 49
E
Ciliary band
Total: 94
Unique: 33
51
23
Apical plate
cells
Total: 250
Unique:
138
13
74
Neuronal
cells
Total: 158
Unique: 74
Upper oral
ectoderm
Total: 122
Unique: 68
30
35
Lower
oral ectoderm
Total: 83
Unique: 43
11
21
Aboral
ectoderm
Total: 107
Unique: 62

Figure 3.24. Regulatory states of the 3 dpf *S. purpuratus* larva. A) Comparison of the transcription factor content per germ layer. Venn diagram indicating the shared and unique transcription factors per germ layer. Ectodermally derived cell types are shown in blue, mesodermally derived in red, and endodermally derived in yellow. B) Comparison of the transcription factor content across mesodermal lineage. Venn diagram showing the shared and unique transcription factors per comparison. C) Transcription factor content comparison of the pyloric sphincter (endodermally derived) and skeletal cells (mesodermally derived), used as a negative control of our comparison. D) Comparison of the transcription factor content per endodermal lineage and endodermally derived cell types. Venn diagram showing the shared and unique transcription factors per comparison. E) TF signature comparison of ectodermally derived cell types. Venn diagram showing the shared and unique transcription factors per comparison. Cartoons indicated the relative position of each cell type/lineage. Mesodermal cell types/lineages are shown in shades of red, endodermal ones in shades of yellow and endodermal ones in shades of blue.

Another observation that arises from this analysis is that neighboring cell types, and in most cases, cell types with common developmental origins, share a larger number of TFs (Fig. 3.24 D) than cell types specified independently. An example of this is the skeletal and pyloric sphincter cells that have independent developmental histories and seem to share only seven transcription factors (Fig. 3.24 C).

The larval cell types' regulatory profile was further analyzed by identifying the expression domains of members of significant TF families (Fig. 3.25). The *S. purpuratus* Homeobox transcription factor family was amongst the first to be identified in this species. Howard-Ashby and co-authors demonstrated that most of them are already expressed by the gastrula stage (2 dpf) and that several Hox family members are expressed in domains deriving from all three germ layers (Howard-Ashby et al., 2006). The 3dpf single cell analysis supports the same notion by providing missing information on the Homeobox TF expression's compartmentalization to different cell types. Plotting of all the Hox family transcription factors detected in the early pluteus dataset revealed that most of the Homeobox classes are enriched in ectodermally derived cell types such as the apical plate and neurons, whereas members of the HOXL subclass of the ANTP Class and the HNF class appear to be enriched in the endodermally derived cell types (Fig. 3.25). Other prominent transcription factor families, including the Forkhead and Ets, shown to be expressed throughout sea urchin embryogenesis (Rizzo et al. 2006, Tu et al. 2006), also appear to be expressed in a broad cell type spectrum (Fig. 3.25). In more detail, Forkhead transcription factors are highly expressed in specific cell types of all three germ layer derivatives (Fig. 3.25), whereas the Ets family TFs are enriched in the ectodermally and mesodermally derived cell types (Fig. 3.25).

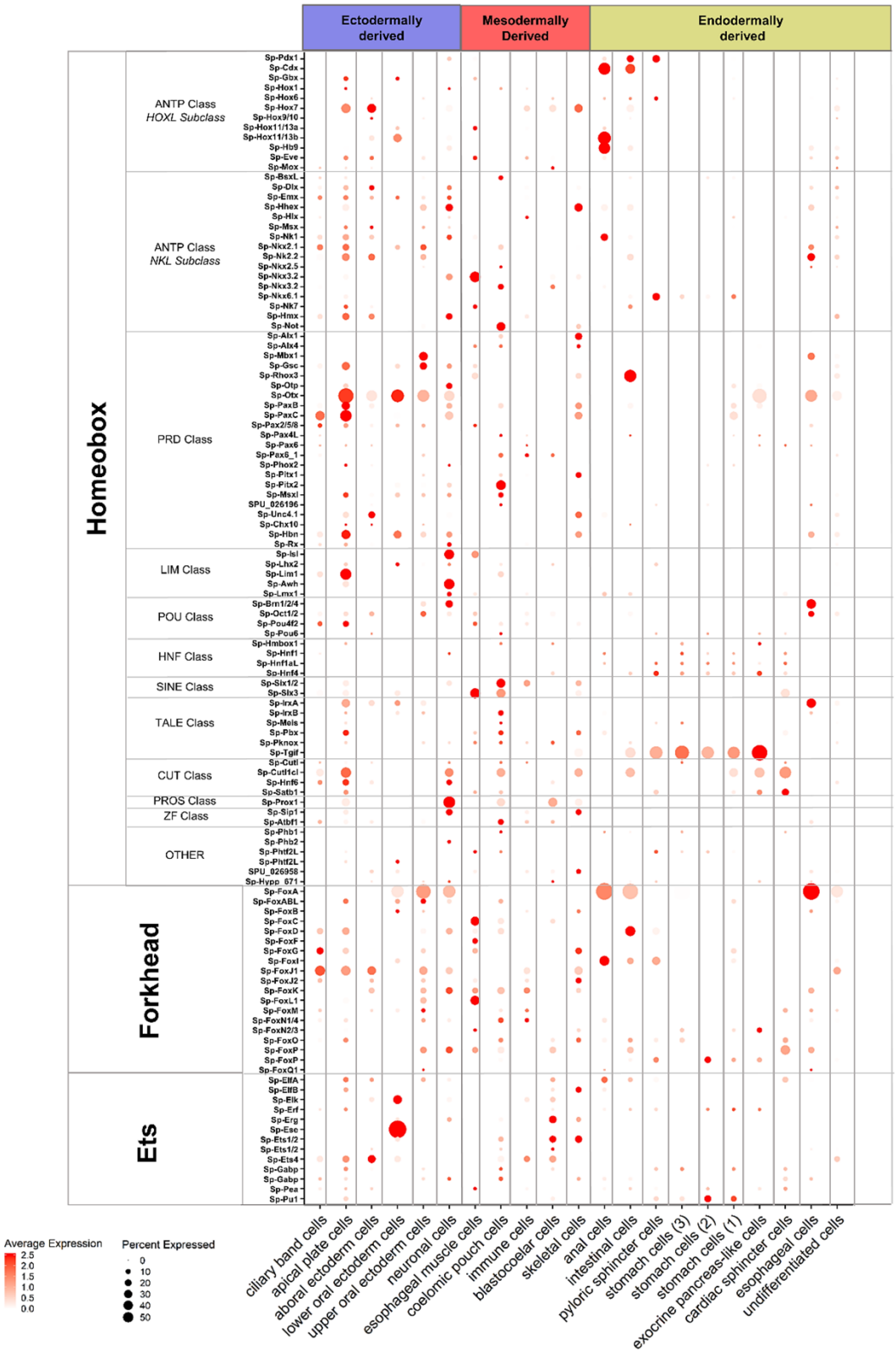


Figure 3.25. Expression domains of the transcription factor family members. Dotplot showing the average expression of members of the Homeobox, Forkhead, and Ets transcription factor families. Each cell type's developmental origins are color-coded, blue for ectodermally derived, red for mesodermally derived, and yellow for endodermally derived ones.

The active regulatory state of a given cell type is an immediate consequence of the GRN wiring that cell type at a given time point. Over decades of sea urchin research, a significant amount of GRNs has been disentangled, focusing mostly on the TFs that are an essential part of them. In this PhD work, two already described gene regulatory modules and one GRN were used to show how scRNA-seq can be used to identify gene regulatory modules and possible novel domains of expression and thus novel function.

The first gene regulatory module tested was identified by Martik et al., and it was shown to control the homing of the small micromeres to the coelomic pouches while being similar to the gene regulatory module controlling retinal specification in the fruit fly (Martik and McClay 2015). Using the single-cell atlas to plot the genes that constitute this module resulted in their identification in the coelomic pouch cell type as previously demonstrated, as well as to a novel domain: the apical plate, suggesting that the same module might be operating in this additional domain (Fig. 3.26 A). Similarly, when plotting genes wiring the aboral ectoderm gene regulatory module (Ben-Tabou de-Leon et al. 2013), all of the genes were present in the aboral ectoderm broad cell type as well as in the apical plate one (Fig. 3.26 B).

In the case of the skeletogenic GRN, drafted at pre-gastrula stages, which is amongst the most well-characterized GRNs of a sea urchin differentiated cell type, plotting of its gene members (Ettensohn 2020) revealed that most of its components are still active at the pluteus stage and differentially expressed in the skeletal broad cell type (Fig. 3.26 C).

Finally, gaining confidence that the 3 dpf scRNA-seq dataset accurately predicts GRN wirings and complex molecular fingerprints, the single cell information was used to recreate a previously published endodermal gene expression atlas reconstructed at the same developmental stage (Annunziata et al., 2014). Strikingly, the scRNA-seq based atlas is almost identical to the one generated via applying traditional molecular biology techniques while simultaneously adding more information on the average expression and the percentage of cells expressing those gene markers (Fig. 3.27).

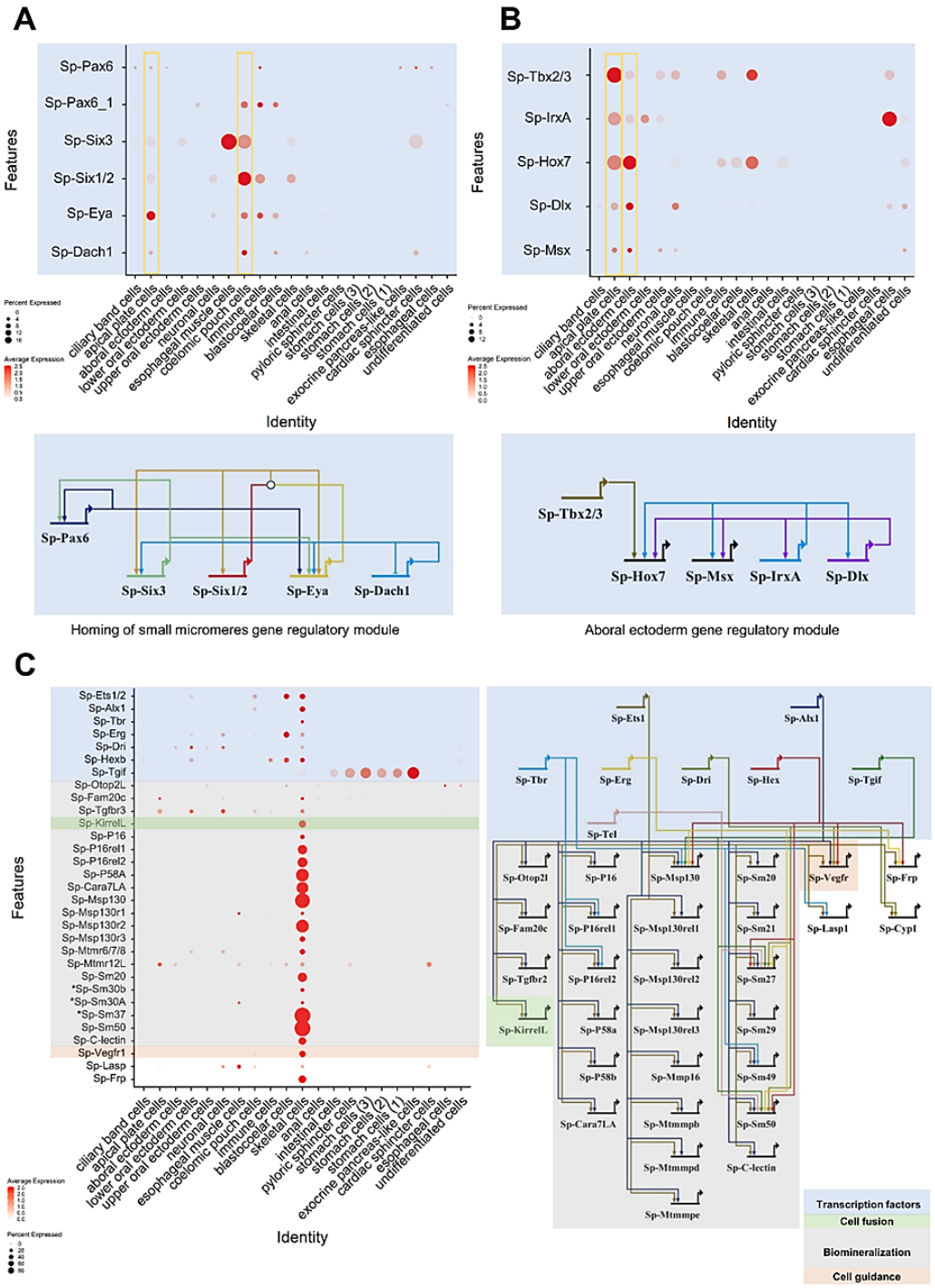


Figure 3.26. Validation of preexisting GRNs and putative novel function of specific gene regulatory modules. A) Dotplot showing the average expression of genes involved in the homing of the small micromeres to the coelomic pouch gene regulatory module. The average expression of genes involved in this module is detected in the coelomic pouch putative cell type as previously described and in a novel domain: the apical plate. B) Dotplot of the expression pattern of genes involved in an aboral ectoderm gene regulatory module. All members are present in the aboral ectoderm putative cell type as well as the apical plate one. C) Pre-gastrula gene regulatory network. Asterisks indicate larval genes involved in biomineralization, putative members of this GRN.

The inconsistency of that very first endodermal cells atlas with the scRNA-seq one derives from the fact that the *in situ* hybridization atlas was based on the investigated genes' topological expression not taking into account that several endodermal domains harbor more than once cell types. For instance, the mRNAs for the transcription factor *SoxC* were detected by *in situ* hybridization in the esophageal region, and thus *SoxC* was considered an esophageal domain molecular marker. ScRNA-seq shows that although SoxC is expressed in the esophageal region, it is expressed in cells adopting a neuronal fate, rather than an esophageal one.

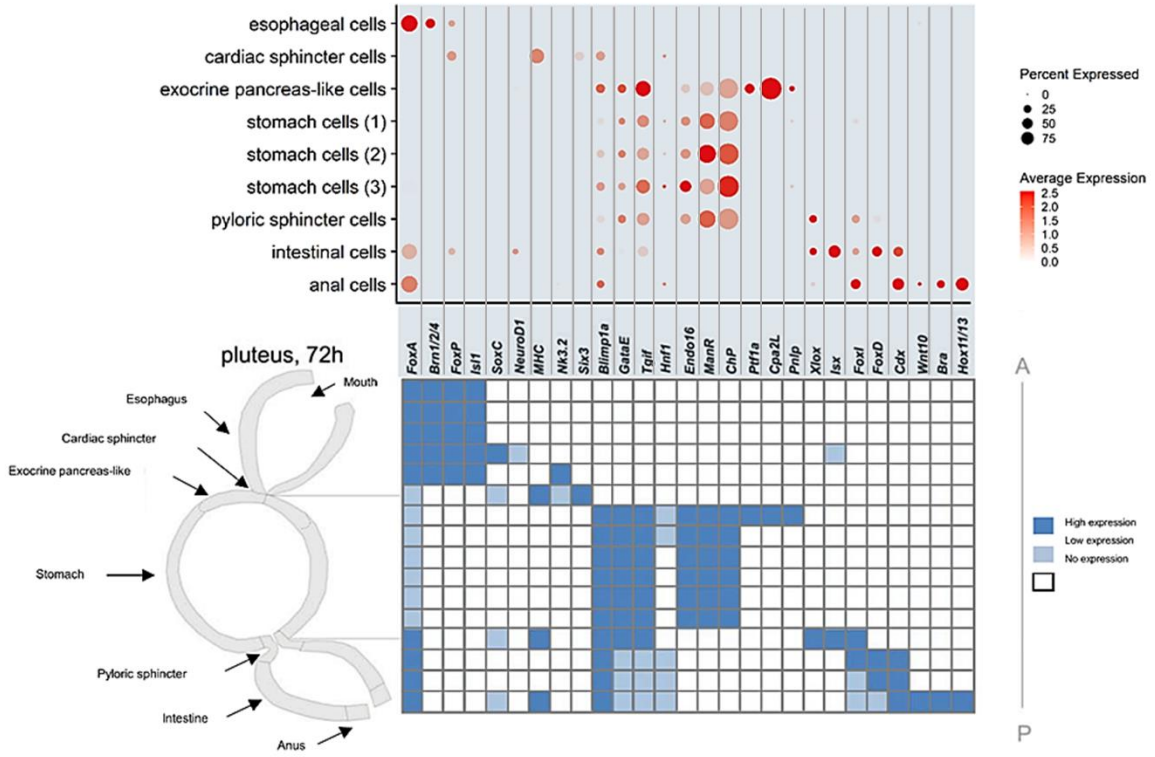


Figure 3.27. Endodermal cell types at a single-cell resolution. Dotplot showing the average expression of marker genes labeling specific endodermal regions. scRNA-seq prediction recapitulates the previously described compartmentalization of gene expression along the 3 dpf sea urchin larva's digestive tract. (Heatmap is adapted from Annunziata et al. 2014).

Chapter 4

Neuronal diversity during sea urchin development

In this chapter, I continue my analysis to investigate the extent of neuronal diversity during sea urchin development, as indicated by the scRNA-seq data. After I successfully recognized the cell type content of the developmental time-points of interest, I set out to provide a thorough characterization of the embryonic and larval nervous system. This chapter contains the results of this analysis ranging from a general characterization of the neuronal types and identity to describing specialized neuronal types. Detailed results discussed in this chapter can be found in the Non-book component of the thesis.

4.1. Untangling neuronal diversity

Nervous systems control all aspects of animal development and response to environmental stimuli, and their organization is a crucial feature of animal body plans. Diverse neurons regulate distinct processes, and their function results in the generation of specific behaviors.

Independently of their identity, all neurons arise from neural progenitors and stem cells through neurogenesis. Neurogenesis occurs during embryonic development in domains with neurogenic potentials in which the developmental program of neural progenitor cells diverges to give rise to a distinct neuronal type. Sea urchin neurogenesis has been described at a molecular level, and it is evident that sea urchin neurogenesis uses a bilaterian conserved “toolkit” (Hartenstein and Stollewerk 2015).

This toolkit consists of a generalized expression of proneural transcription factors controlled by Delta-Notch lateral inhibition (e.g., *SoxB*) that, due to restrictive mechanisms (Wnt and Tgf-β signalings), is limited to specialized neuroectodermal domains (anterior neural ectoderm, ANE, and ciliary band). Those ectodermal cells are destined to directly differentiate to epithelial sensory neurons and ganglia, with the ANE cells producing more neurons.

In most cases, neuronal communication is based on an electric current. The action potential, produced by one neuron, travels along the axon towards the nerve terminal, where it causes calcium-dependent fusion of synaptic vesicles containing neuromodulators with the plasma membrane causing their release in the synapsis. Neuromodulators (neurotransmitters and neuropeptides) bind to their receptors expressed by the target cells and initiate an intercellular signaling response. This form of chemical signaling arose during the earliest evolution of living forms as a means of cell communication. Neuromodulators are present in all organisms studied so far, including organisms that lack a nervous system. For instance, neuromodulators have been found in unicellular organisms with a chemotactic/chemorepellant role in plants as well as sponges and placozoans (reviewed in D’Aniello et al. 2020). The presence of the neuropeptidergic/neurotransmitter signaling in taxa without a nervous system suggests that this endocrine system predates the formation of the nervous system. Moreover, it highlights its dominant role in cell communication since

neurotransmitters and receptors had to go through strong evolutionary forces and evolved parallel to retain the signaling transduction ability (Hoyle 2011).

Most nervous system components have been classified based on the combination of different neuromodulators that a neuron produces. Thus, the understanding of the neuromodulators a neuron produces contributes to identifying its function.

The sea urchin nervous system has been among the first echinoderm cell types to be characterized at a molecular level, yet the mechanisms that give rise to diverse neuronal types and their number remain partially unknown.

The echinoderm nervous system organization is a distinguishing feature among other deuterostomes. Both the adult and the larva sea urchin have nervous systems thought to be completely independent (Burke et al. 2006). The echinoderm nervous systems are decentralized, but they are not a simple nerve net since the radial nerves are segmentally organized. Extensive work has been carried out to identify the signaling cascades that guide neuronal specification and the subsequent events that occur during neuronal differentiation. However, this is limited to describing broad neuronal classes based on the detection of neuropeptides, neurotransmitters, and enzymes involved in their biosynthesis. Such studies revealed that the sea urchin larva consists of neurons originating from the ectoderm placed within the apical domain, or scattered throughout the ciliary band, as well as close to it, and endodermally derived neurons that pattern the esophageal region, sphincters, and control muscle contraction (Fig. 4.1).

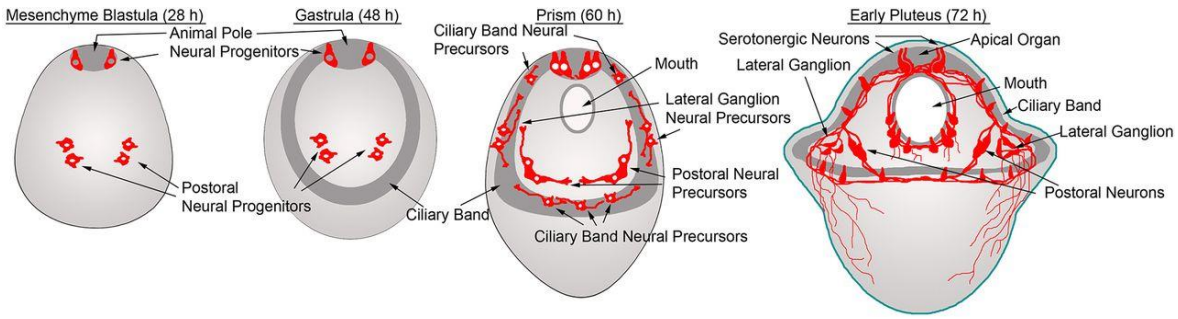


Figure 4.1. The stages of neurogenesis in sea urchin embryos. Neural progenitors are present at blastula (1 dpf) stage in the anterior neuroectoderm and in cells where the post-oral arms will form. These progenitors will give rise to neural precursors in the same domains at gastrula (2 dpf), while precursors for ciliary band neurons arise in the ciliary band or the oral and aboral ectoderm near the ciliary band. At pluteus stage (3 dpf), the nervous system is formed, and extensive neuronal projections are evident all around the ciliary band and the apical plate of the larva (Garner et al. 2016).

Previous genome-based studies identified genes encoding for transcription factors and signalings involved in neurogenesis, neuromodulators, and synaptic release machinery associated genes, providing a broad view of the sea urchin nervous system. The majority of information regarding sea urchin neurogenesis and neuronal diversity comes from studies carried out in various sea urchin species, developing the bias that the same neuronal types are present across sea urchins. A recent study by Massri et al., in which the authors identified gene expression heterochronies between the two sea urchin species *S. purpuratus* and *Lytechinus variegatus,* that vary in magnitude and are present throughout development in all germ layer derivatives (Massri et al. 2020). This observation supports the possibility of nervous system diversification across sea urchin species since heterochronic events, if not buffered at a gene regulatory level, could lead to diverse specification pathways and different neuronal types.

The most recent information on *S. purpuratus* embryonic and larval neuronal diversity comes from Wood et al., 2018. The authors used co-expression analyses of novel neuromodulators and identified two distinct neuronal types at 2 dpf gastrula and seven at 3 dpf pluteus developmental stages. In more detail, the authors found that the two neuronal types identified at gastrula stage are spatially associated with the apical plate and oral ectoderm. In contrast, regarding the pluteus stage, they found an apical organ, an oral distal, a post-oral, a lateral ganglia, two associated with the mouth, and one midgut neuronal populations (Fig. 4.2).

In this chapter of the PhD thesis, I focus my attention on investigating the embryonic and larval nervous system organization by using the already published and available information combined with targeted analyses of the neuronal associated putative broad cell types.

Despite the report of various neuronal populations present at gastrula and early pluteus stages, the initial clustering analysis and annotation based on publicly available neuronal gene markers concluded the presence of a single neuronal broad cell type (Fig. 4.3). This suggests that either the different neuronal types described so far are mostly transcriptionally similar or that the number of cells captured is insufficient to reveal these differences. Furthermore, integration of the 2 and 3 dpf datasets, which leads to a higher number of analyzed cells, still produces only one neuronal putative cell type (Fig. 4.3).

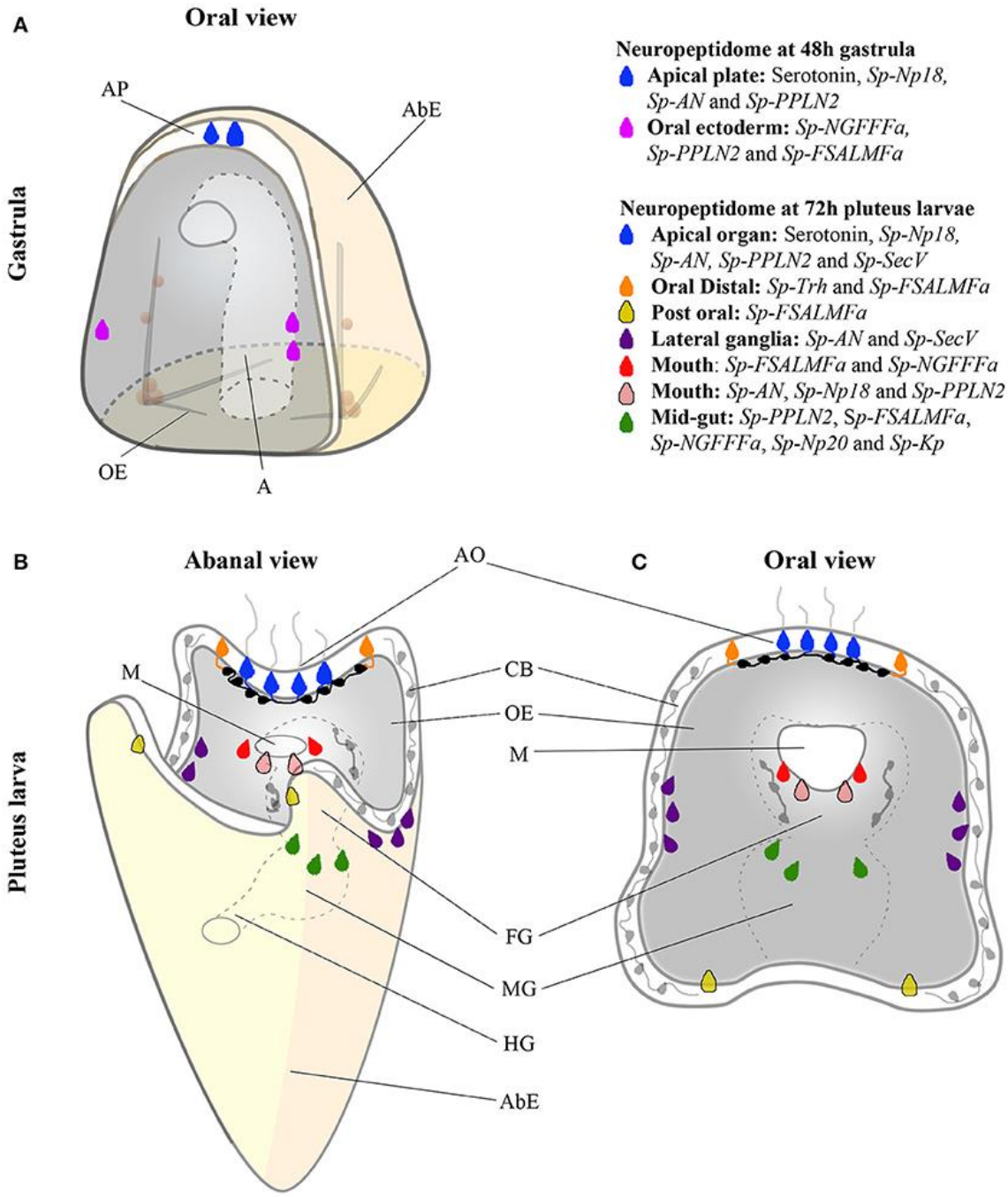


Figure 4.2. Neuronal types present at 2 and 3 dpf of *S. purpuratus* development. A) Schematic representation of 2 dpf gastrula embryo showing the different neuronal types present based on their neuromodulator content. B) Schematic representation of a 3 dpf pluteus larva from a ventral lateral point of view. C) Schematic representation of a 3dpf pluteus larva from an oral point of view. AP (Apical Plate), AbE (Aboral Ectoderm), OE (Oral Ectoderm), A (Archenteron), AO (Apical Organ), CB (Ciliary band), M (Mouth), FG (Foregut), MG (Midgut), and HG (Hindgut) (Figure from Wood et al. 2018).

It has been previously demonstrated that that neuronal specification and several differentiation steps occur through a similar gene regulatory toolkit in *S. purpuratus* and *L. variegatus* sea urchin species (Figs. 4.5 and 4.6). Taken together, these data favor the hypothesis that the differentiating/differentiated neuronal types are

transcriptionally very similar, explaining the presence of only one neuronal cluster at gastrula and early pluteus stages.

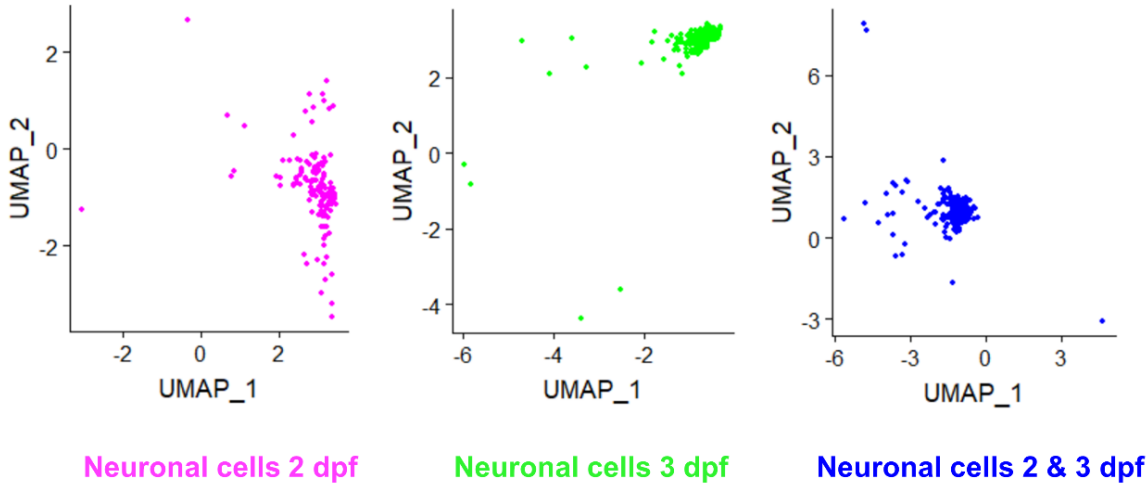


Figure 4.3. Neuronal putative broad cell type. UMAP of the single neuronal putative broad cell type at 2 dpf gastrula (magenta) and 3 dpf pluteus (green) stages. The integration of the two datasets still results in the formation of one neuronal cluster (blue) and a similar UMAP coming from the integration of the 2 and 3 dpf datasets.

To investigate this, subclustering and re-analysis of the neuronal broad putative cell types at gastrula and early pluteus time points were performed. To maximize the amount of information we could retrieve from this analysis, a resolution of 1.9 was chosen, which produces the highest number of clusters, while at the same time, each cluster consists of five or more cells. This resulted in the subclustering of the single neuronal cluster into 6 gastrula and 12 early pluteus neuronal populations (Fig. 4.4).

Next, in order to identify the neuronal types revealed by the subclustering analysis, I took advantage of the existing information on sea urchin neurogenesis and neuronal differentiation (Burke et al. 2006, Wei et al. 2009, Garner et al. 2016, Wei et al. 2016, McClay et al. 2018, Perillo et al. 2018, Wood et al. 2018) and selected known gene markers in an attempt to identify the different neuronal populations. Plotting genes involved in sea urchin neurogenesis, neuronal differentiation, neurotransmitters, and neuropeptides that hint distinct neuronal function allowed the identification of the molecular signature of the distinct neuronal subtypes and the recognition of their spatial location on the embryo and larva.

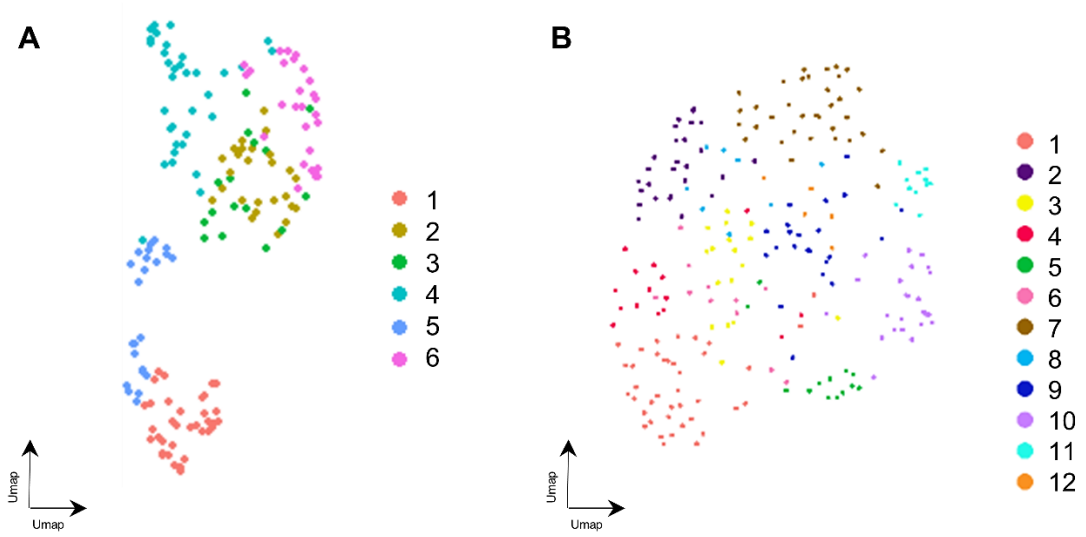


Figure 4.4. Subclustering analysis reveals multiple neuronal populations. A) The single neuronal cluster from the 2 dpf gastrula dataset contains 6 transcriptionally distinct populations. B) The single neuronal cluster from the 3 dpf pluteus dataset contains 12 transcriptionally distinct populations.

The sea urchin neuronal differentiation pathway has been described in great detail, and it has been shown to go through a stepwise differentiation program (Fig. 4.5) that includes the transient expression of the Notch ligand Delta followed by the expression of the transcription factors SoxC and Brn1/2/4 (Garner et al., 2016).

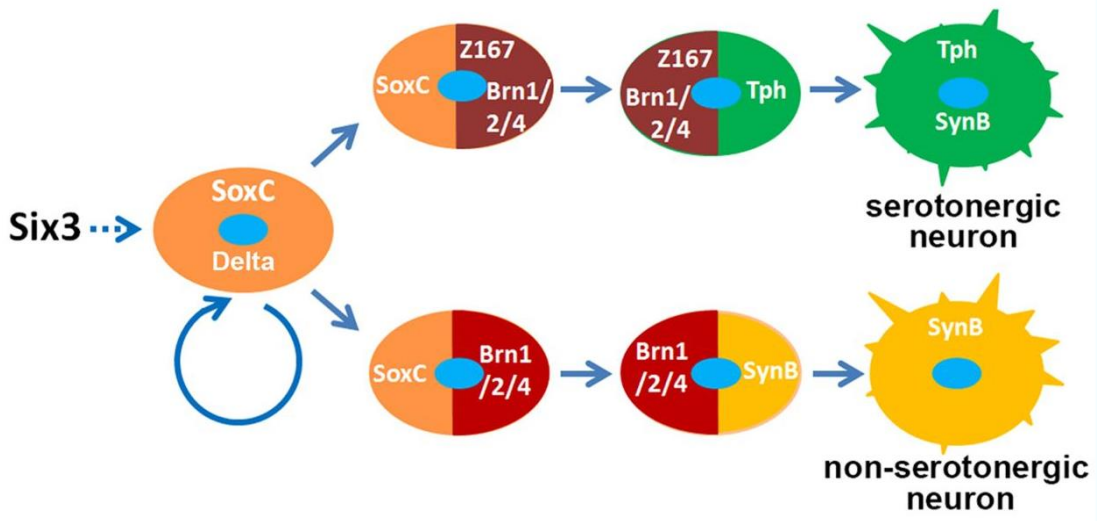


Figure 4.5. Neuronal differentiation cascade in sea urchin. The transcription factor *SoxC* is activated in isolated neural precursors, which are dividing (circular arrow), and some start to express the Notch ligand Delta and are subject to lateral inhibition. *SoxC* is required for the development of all neuronal types, while the expression of the transcription factor *Brn1/2/4* depends on *SoxC*. All developing neurons express *Brn1/2/4* while developing serotonergic (Tph positive) neurons also express the transcription factor *Z167* (Wei et al. 2016).

The transcription factors *Sip1, Z167, Ngn,* and *Otp* are shown to be involved in the final steps of the differentiation pathway (Fig. 4.6) that will give rise to diverse neuronal populations, including apical and ciliary band neurons (Wei et al. 2016, McClay et al. 2018). The more in-depth characterized neuronal population is the apical organ serotonergic neurons that are already in place at 2 dpf gastrula and appear as 2-4 bilateral pairs of cells.

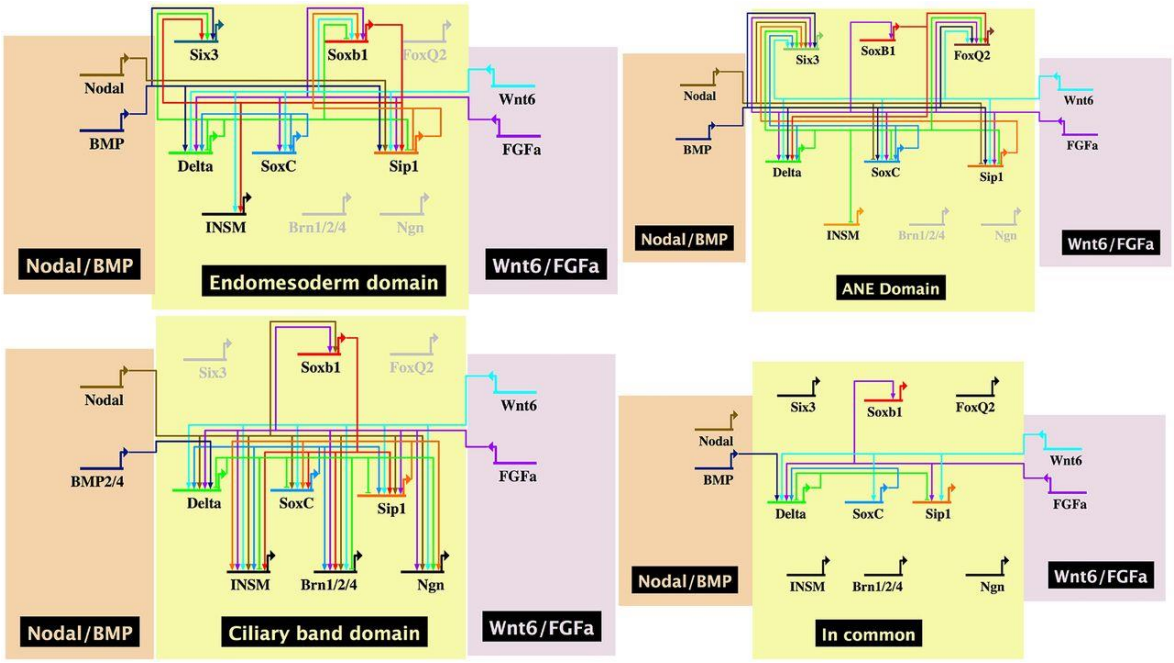


Figure 4.6. GRNs guiding neurogenesis in the sea urchin *L. variegatus*. Neuronal specification GRNs and their common components guiding the specification of neuronal progenitors at the three neurogenic domains of the embryo (McClay et al. 2018).

The 2 dpf gastrula dataset, subclustering analysis revealed 6 distinct neuronal types deriving from the single neuronal putative cell type. Plotting of the known gene markers and performing *in situ* hybridization and IHC experiments allowed the identification of the neuronal clusters and mapping them back on distinct embryonic domains.

According to the average expression of the *Sp-Delta*, *Sp-SoxC,* and *Sp-Brn1/2/4,* it seems that one neuronal population corresponds to developing neurons that are undergoing differentiation (Fig. 4.7A). FISH of *Sp-Brn1/2/4* and *Sp-SoxC* (Fig. 4.7 B3, B5) revealed that most of the *Sp-SoxC* and *Sp-Brn1/2/4* cells are in the oral ectoderm and endodermal domains suggesting the developing neurons population also reside in those domains. On the other hand, the average expression of the transcription factors *Sp-Nkx2.1*, *Sp-Fez*, *Sp-Z167*, *Sp-Hbn,* and *Sp-Rx* suggests

that the rest of the populations that arise from the subclustering analysis are associated with the anterior neuroectoderm region (Fig. 4.7A).

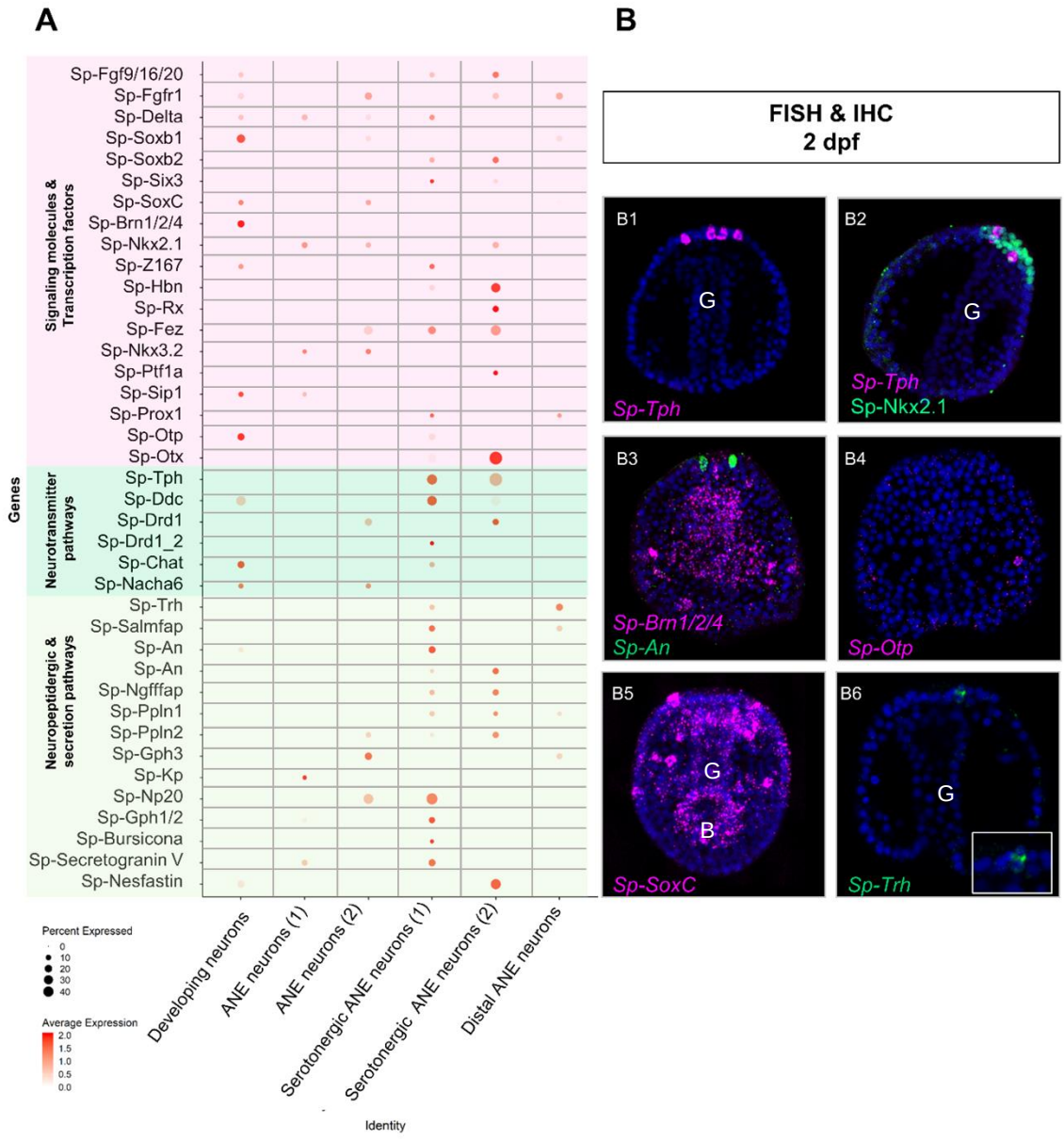


Figure 4.7. Neuronal complexity of the 2 dpf *S. purpuratus* embryo. A) Dotplot showing the average expression of signaling molecules and transcription factors involved in sea urchin neurogenesis/neuronal differentiation as well as neurotransmitter and neuropeptidergic pathways. B) FISH of *S. purpuratus* 2 dpf embryos with specific antisense probes for the neuronal genes *Sp-Tph (*B1-2), *Sp-Brn1/2/4* (B3), *Sp-An* (B3), *Sp-Otp* (B4), *Sp-SoxC* (B5), and *Sp-Trh* (B6). FISH shown in figure B2 is paired with immunohistochemical detection of the transcription factor *Sp-Nkx2.1*. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. B, blastopore; G, gut. *Note: Sp-An appears twice in the plot since two WHL IDs (transcripts) correspond to the same gene*.

Furthermore, the average expression of the rate-limiting enzyme in serotonin biosynthesis (*Sp-Tph*) reveals that two out of the five ANE-associated neurons are serotonergic, produce the neuropeptide AN, and one of them is positive for the transcription factor *Sp-Nkx2.1* (Fig. 4.7A). The predicted spatial distributions *of Sp-Tph, Sp-An, and Sp-Nkx2.1,* are validated by FISH and IHC experiments (Fig. 4.7 B1, B2, and B3).

Sp-Trh is expressed in a neuronal population flanking the larva's apical organ, although its expression pattern was not investigated at 2 dpf gastrula stage. ScRNA-seq data suggest that *Sp-Trh* is indeed expressed, and those predictions were validated by FISH, which revealed *Sp-Trh* transcripts being present in the anterior neuroectoderm region (Fig. 4.7 B6).

Taking advantage of all the available data and the experiments and analysis performed in this thesis allowed the mapping of all the neuronal populations back to specific domains of the embryo (Fig. 4.8).

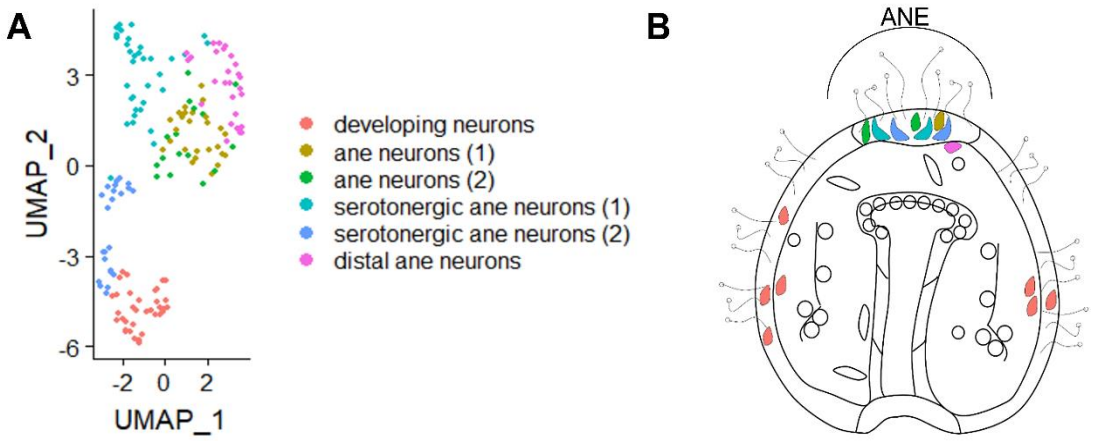


Figure 4.8. Identification of neuronal populations. A) UMAP showing 6 distinct populations as revealed by our subclustering analysis. B) Schematic representation of the 2 dpf gastrula showing the localization of the different neuronal populations. Color-code is consistent with the subclustering analysis. ANE, anterior neuroectoderm

The scRNA-seq approach applied in this thesis allowed the expansion of the number of different neuronal types present at the *S. purpuratus* gastrula stage and increased our understanding of the nervous system complexity as it expands the number of neuronal types present at this stage, from the two described in the Wood et al. 2018 study to six. The six neuronal types I identified agree with the ones identified by Wood et al. 2018 in terms of spatial distribution in two domains: the oral ectoderm

and anterior neuroectoderm. What our approach adds to this, is the identification of the molecular signature complexity of the nervous system by assessing the pool of genes expressed by each neuronal population rather than testing the expression pattern for individual gene candidates that are expressed in more than one neuronal population.

Next, I set out to investigate the neuronal diversity of the 3 dpf early pluteus dataset, and to do so, I applied a similar approach as for the 2 dpf gastrula stage by interrogating the scRNA-seq data for sea urchin neuronal genes and validating the computational predictions (Fig. 4.9) with FISH and IHC (Fig. 4.10).

Sp-Delta, *Sp-SoxC* and *Sp-Brn1/2/4* transcripts co-localize in two neuronal populations suggesting that neuronal differentiation occurs within those two sub-clusters (Fig. 4.9). Interestingly, in one of these populations, *Sp-Rx*, *Sp-Hbn* (Fig. 4.10 J), and *Sp-Six3* (Fig. 4.10 P) transcripts, known to be expressed in the periphery of the larva's apical domain (Burke et al., 2006, Wei et al., 2009), were found. This suggests that this population is located in the apical plate's periphery and not within the apical organ itself. In the same domain, another neuronal type was detected that corresponds to a distal apical plate neuronal population, which is known to be double positive for the *Sp-Trh* and *Sp-Salmfap* neuropeptides (Wood et al. 2018). Additionally, both scRNA-seq and FISH confirm that a third neuropeptide *Sp-Kp* (Kissepeptin) is all expressed in the same neuronal type (Fig. 4.10 Q).

In total, we identified three neuronal populations located in the apical domain of the larva that express *tryptophan hydroxylase* (*Tph*), suggesting that those are the serotonergic neurons of the larva (Figs. 4.9 and 4.10 M), while one also produces the neuropeptide Sp-An (Figs. 4.9 and 4.10 M). The ciliary band is hypothesized to be equivalent to a peripheral nervous system, supporting neurogenesis and harboring sensory neurons (Slota et al., 2020). From this analysis, two distinct cholinergic populations within the ciliary band were found (Fig. 4.9), one of which expresses the enzyme involved in acetylcholine biosynthesis (*Sp-Chat*) and the two nicotinic acetylcholine receptors Sp-*Nacha6* and Sp-*Chrna9* (Fig. 4.10 G, I). Additionally, a neuronal subtype associated with the ciliary band was identified, which neurons are placed close to the ciliary band and corresponds to the lateral and post-oral neurons as judged by the presence of Sp-AN neuropeptide (Wood et al. 2018) (Fig. 4.10 A, B, C, H, N).

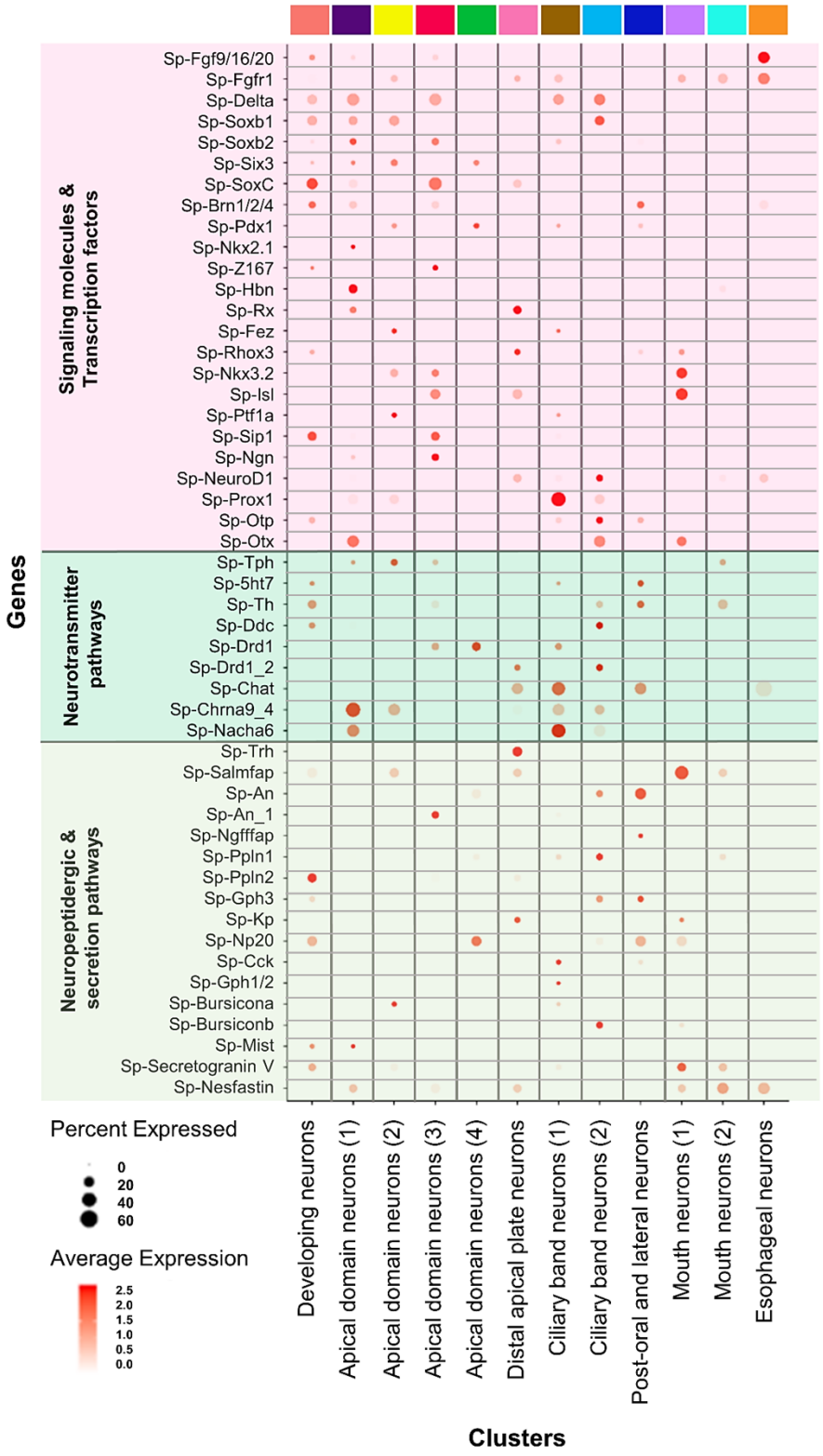


Figure 4.9. Neuronal diversity at 3 dpf *S. purpuratus* pluteus larva. Dotplot showing the average expression of signaling molecules and transcription factors involved in sea urchin neurogenesis/neuronal differentiation as well as neurotransmitter and neuropeptidergic pathways. Color-code is consistent with the subclustering analysis (Fig. 4.4). *Note: Sp-An appears twice in the plot since two WHL IDs (transcripts) correspond to the same gene*.

Using known gene markers that label differentiated neurons expressed in the rim of the larva's mouth such as *Sp-Nkx3.2*, a gene essential for the formation of mouth neurons (Wei et al. 2011), the transcription factor *Sp-Isl* (Perillo et al. 2018), the neuropeptide *Sp-Salmfap* (Wood et al. 2018) and the enzyme *tyrosine hydroxylase* (*Sp-Th*) involved in the dopaminergic pathway, two mouth neurons subtypes (Fig.4.9) were revealed. Finally, one neuronal population, based on its molecular signature, including the genes *Sp-Fgf9/16/20*, *Sp-Fgfr1,* and *Sp-NeuroD1*, seems to be associated with endodermal structures such as the esophagus.

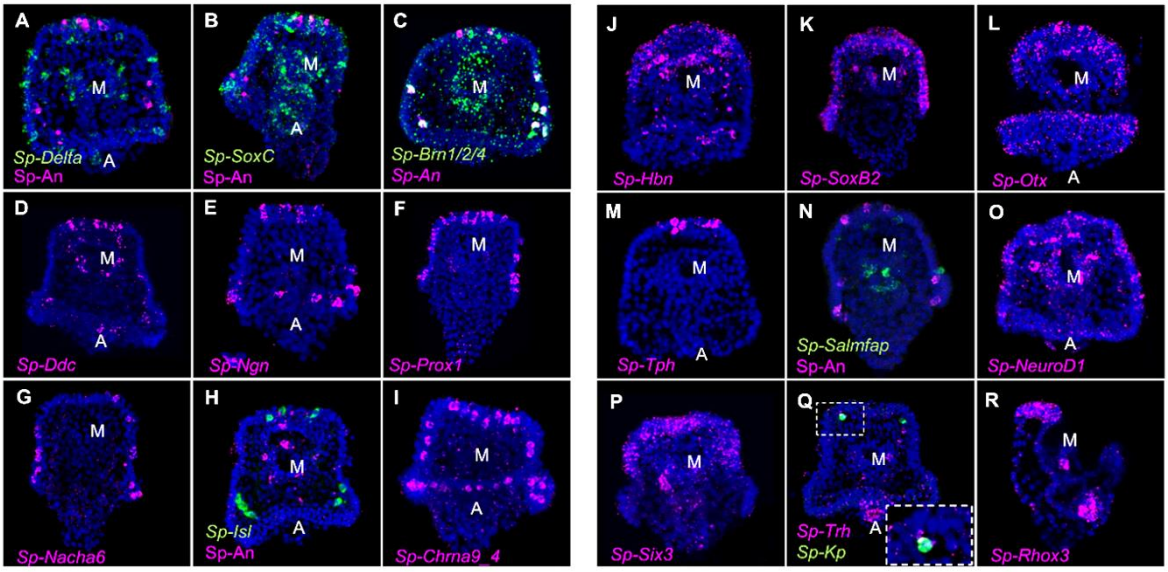


Figure 4.10. Validations of scRNA-seq predictions. FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes for the neuronal genes *Sp-Delta* (A), *Sp-SoxC* (B), *Sp-Brn1/2/4* (C), *Sp-Ddc* (D), *Sp-Ngn* (E), *Sp-Prox1* (F), *Sp-Nacha6* (G), *Sp-Isl* (H), *Sp-An* (C and N), *Sp-Chrna9_4* (I), *Sp-Hbn* (J), *Sp-SoxB2* (K), *Sp-Otx* (L), *Sp-Tph* (M), *Sp-Salmfap* (N), *Sp-NeuroD1*(O), *Sp-Six3* (P), *Sp-Trh* (Q), *Sp-Kp* (Q) and *Sp-Rhox3* (R). FISH shown in figures A, B, and H are paired with immunohistochemical detection of the neuropeptide Sp-An. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. A, anus; M, mouth.

The spatial distribution of all twelve neuronal populations found at this stage can be seen in Figure 4.11, in which the different populations are grouped based on their spatial location on the larva, reflecting their developmental origins from ectoderm and endoderm.

Overall, the subclustering analysis of gastrula and early pluteus time-points increased the resolution of the different neuronal subtypes present at this developmental stage and revealed new neuronal subtypes and novel gene markers as well as gene candidates for future studies and for drafting GRNs.

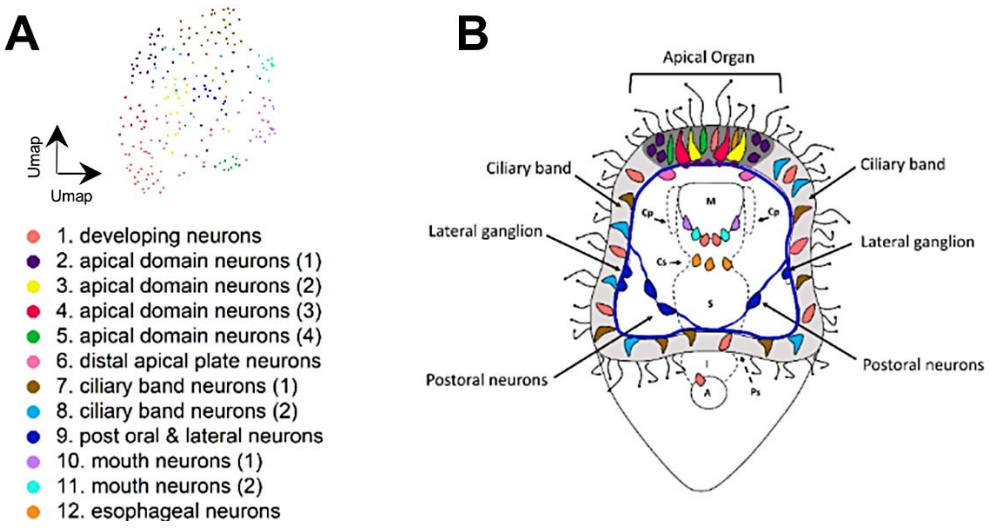


Figure 4.11. Mapping of the larval neuronal populations. A) UMAP showing 12 distinct populations as revealed by our subclustering analysis. B) Schematic representation of the 3 dpf pluteus larva showing the localization of the different neuronal populations. Color-code is consistent with the subclustering analysis (Fig. 4.4.).

4.2. Neuronal signaling

After recognizing the neuronal types present at 2 and 3 dpf developmental stages, I focused my attention on understanding the mechanisms by which the nervous system functions. Neurotransmitters are neuromodulators released by nerve cells that initiate intercellular signaling cascades, causing a physiological or behavioral response. According to the release mechanism employed, neurotransmitters are grouped into two categories: conventional and non-conventional. Conventional neurotransmitters are contained in synaptic vesicles and released from the presynaptic nerve terminal with exocytosis. Conversely, unconventional neurotransmitters are not contained in vesicles and often function by initiating signaling from the postsynaptic to the presynaptic cells.

Conventional neurotransmitters are usually small-molecule neurotransmitters that span from individual amino acids such as GABA and glutamate to small peptides and biogenic amines. Known biogenic amines neurotransmitters are catecholamines (dopamine, epinephrine, and norepinephrine), the indolamine serotonin, and the imidazoleamine histamine. On the other hand, endocannabinoids and nitric oxide belong to the non-conventional group of neurotransmitters, and their neuronal role has been well established in a variety of model systems.

As previously mentioned, the initial characterization of the sea urchin nervous system was solely based on identifying the neurotransmitters that different neuronal populations produce either by *in vivo* visualization methods or by pharmacological inhibition of their signalings. This led to a very well characterized map of neurotransmitter expression and a generalized understanding of their role in development and signal transduction. However, little is known about the mechanisms by which sea urchin neurotransmitters function and which other neuronal signaling components are in place.

In this part of the thesis, I characterize the expression domains of the main conventional and non-conventional neurotransmitter signaling components that are in place at the early *S. purpuratus* pluteus stage.

The majority of the available literature on sea urchin neuronal signaling is focused on the neuronal effect on non-neuronal cell types. To perform an analysis that is compatible with the available observations and check whether these observations are verified at a single cell transcriptomic level, I focus my analysis on the general clustering analysis of the 3 dpf pluteus larva, in which all the putative broad cell types are present. Such an approach allowed me to investigate how the nervous system functions with respect to other cell types rather than understanding the interneuronal signalings in place.

4.2.1. Serotonin signaling

Serotonin [5-hydroxytryptamine (5-HT)] is a neurotransmitter present in the nervous systems of most metazoans (Hay-Schmidt 2000). In vertebrates, serotonin is involved in the regulation of many processes, such as sleeping and feeding and behaviors regarding mood. In echinoderms, serotonin signaling is necessary for normal embryonic development as blocking of the serotonin receptors with antagonists and agonists results in embryonic malformations. The development of such embryos can be restored upon the provision of reverse drugs or serotonin itself, suggesting that the malformations are specifically due to serotonin signaling. (Buznikov et al., 2005). Furthermore, it has been demonstrated that in the sea urchin *Hemicentrotus pulcherrimus*, inhibition of serotonin synthesis reduces larval swimming (Yaguchi and Katow 2003, Katow et al. 2007). Genomic characterization of the *S. purpuratus* serotonin signaling resulted in identifying the enzymes involved

in the serotonin production, putative receptors through which serotonin initiates its signalings to cell type targets.

Plotting of genes that are involved in the biosynthesis of serotonin verified the already published data that the enzymes involved in serotonin biosynthesis are mostly found in the neuronal putative broad cell type (Fig. 4.12). Furthermore, doing the same for serotonin receptors led to the realization that serotonin can potentially initiate a signaling cascade in almost all the putative cell types (Fig. 4.12). The possible multifunctional role of serotonin, as revealed by the scRNA-seq analysis, is in complete agreement with the phenotype obtained when serotonin signaling was blocked in developing sea urchin embryos. Moreover, at least two serotonin receptors appear to be present in the ciliary band putative cell type suggesting that the serotonin signaling system might regulate ciliary beating and swimming ability in a similar way as demonstrated in *H. pulcherrimus.* On the other hand, when interrogating the scRNA-seq data for enzymes such as monoamine oxidase (MOA) and catechol-O-methyltransferase (COMT), that are involved in serotonin and dopamine inactivation/degradation in the synaptic cleft and pre-synaptic terminal, I found an enrichment of such enzymes in mesodermally derivate cell types (Fig. 4.12). In more detail, such genes were enriched in the immune, blastocoelar, and skeletal cells. The common feature of these cells is that they inhabit the blastocoelar space of the larva, while immune and blastocoelar cells are motile cell populations. This means that they are in contact with the internal space and extracellular matrix of the larva in which the neurotransmitters are released. Taken together, this suggests that these neighboring mesodermal cell types might be able to regulate the amount of serotonin and dopamine available in the synaptic cleft by uptaking and degrading them.

Melatonin mediated signaling is evolutionary widespread and has been reutilized throughout evolution to control different physiological responses, with the most well-known to be the regulation of circadian rhythm (Zhao et al. 2019). Melatonin derives from serotonin, which is converted to N-acetylserotonin by serotonin N-acetyltransferase and subsequently transformed to melatonin in a reaction catalyzed by acetylserotonin O-methyltransferase.

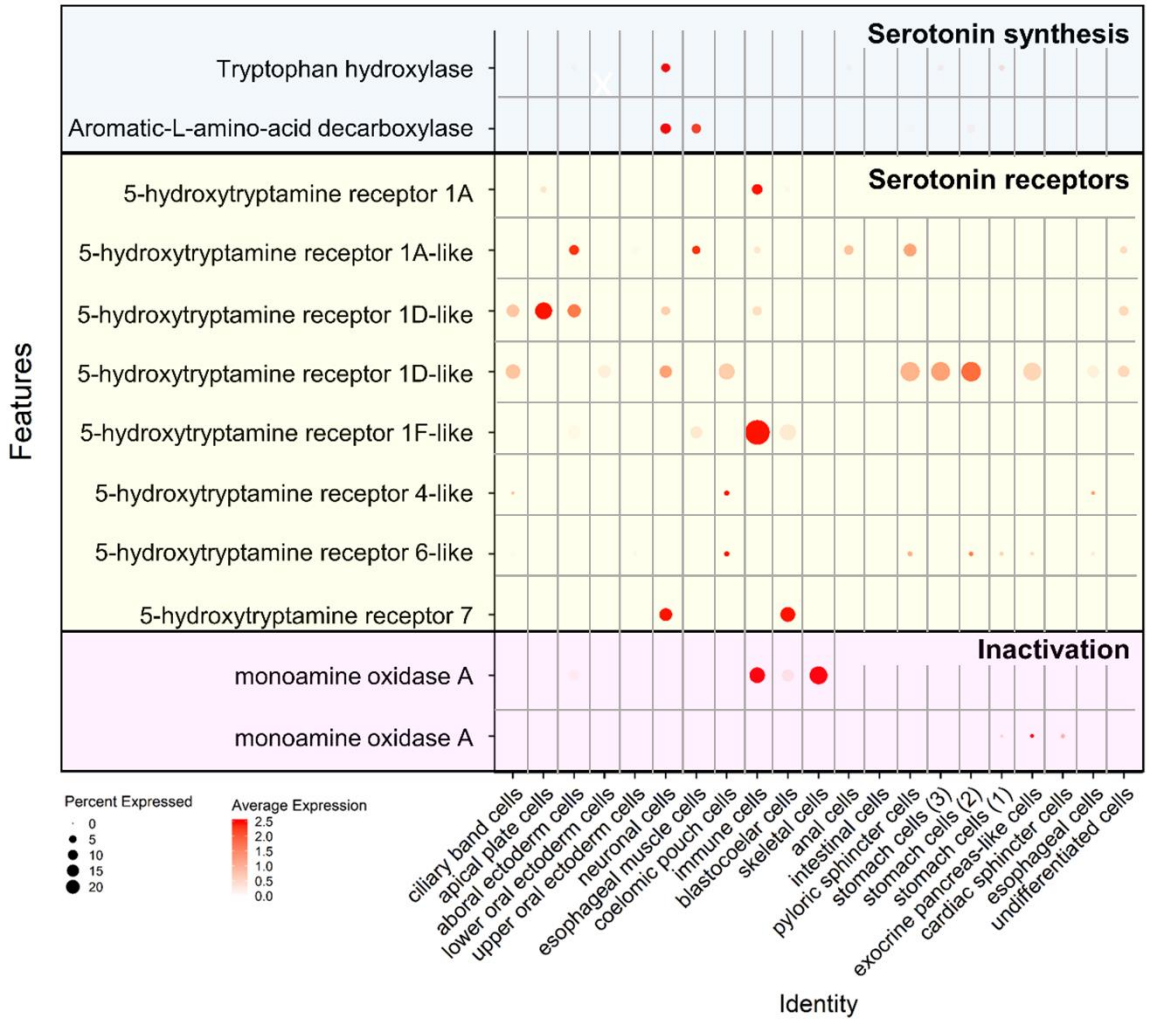


Figure 4.12. Serotonin signaling at 3 dpf larva. Dotplot showing the average expression of genes involved in the biosynthesis, inactivation, and signaling cascade of serotonin.

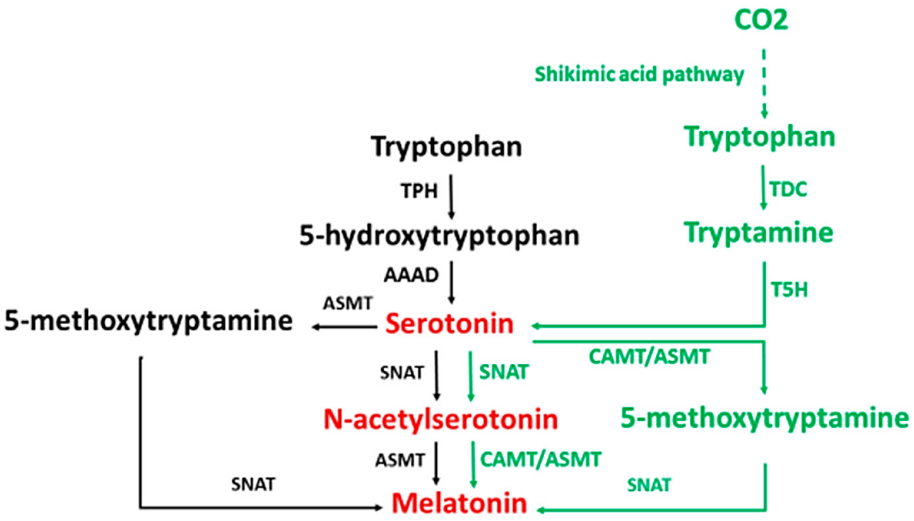


Figure 4.13. Melatonin biosynthetic pathway. Serotonin is being catabolized to N-acetylserotonin and melatonin by the enzymes serotonin N-acetyltransferase (SNAT) and acetylserotonin O-methyltransferase ASMT (Tan et al. 2015).

In a study by Burke et al., the authors provided a detailed characterization of the *S. purpuratus* nervous system based on genomic analysis, no putative genes encoding for serotonin N-acetyltransferase and acetylserotonin O-methyltransferase were found (Burke et al. 2006). These genes are essential for the production of melatonin from serotonin, and since neither gene was found in either ascidians or sea urchins, it was claimed that the production of melatonin is a vertebrate feature. Surprisingly, a recent study by Ding et al. identified that melatonin is present in the sea cucumber *Apostichopus japonicus* and found that melatonin exhibits a sedative effect on locomotion (Ding et al. 2019). Moreover, Squires and colleagues demonstrated a putative mechanism on the metabolism of serotonin in *S. purpuratus* and proposed that N-acetyltransferases encoded in the genome can produce N-acetylserotonin (Squires et al. 2010).

Furthermore, when the recently released version 5.0 of the *S. purpuratus* genome was interrogated, it revealed that the N-acetylserotonin-O-methyltransferase (ASMT) catechol-O-methyltransferase (COMT) genes are encoded in it. Plotting for the genes encoding for the enzymes involved in the melatonin biosynthetic and signaling pathway revealed a broad expression pattern of those genes across cell types (Fig. 4.14).

Nonetheless, only the cell types in which the combination of either or both of the Catechol O-methyltransferase and N-acetylserotonin-O-methyltransferase and N-acetyltransferases is present. As a result, only the ciliary band, immune, blastocoelar, anal, and pyloric sphincter cells can putatively produce melatonin, provided that serotonin is present in them. In addition to the biosynthetic enzyme, transcripts for putative melatonin receptors were found enriched in the immune, blastocoelar, skeletal, and apical plate cell types, suggesting putative autocrine signaling of melatonin to the cells that are producing it (Fig. 4.14). Furthermore, since both immune and blastocoelar cells are part of the sea urchin larva's immune defense system, these data imply that melatonin could be involved in the immune response. Moreover, as previously mentioned, these two cell types consist of cells with increased migratory ability throughout the blastocoel and, in the case of immune cells, within the ectodermal epithelium. The presence of melatonin receptors in these cell types raises the question of whether melatonin signaling is involved in the regulation of locomotion as in the sea cucumber. Future studies are

needed to show whether melatonin binds to those receptors as well as to clarify the function of the melatonin induced signaling response.

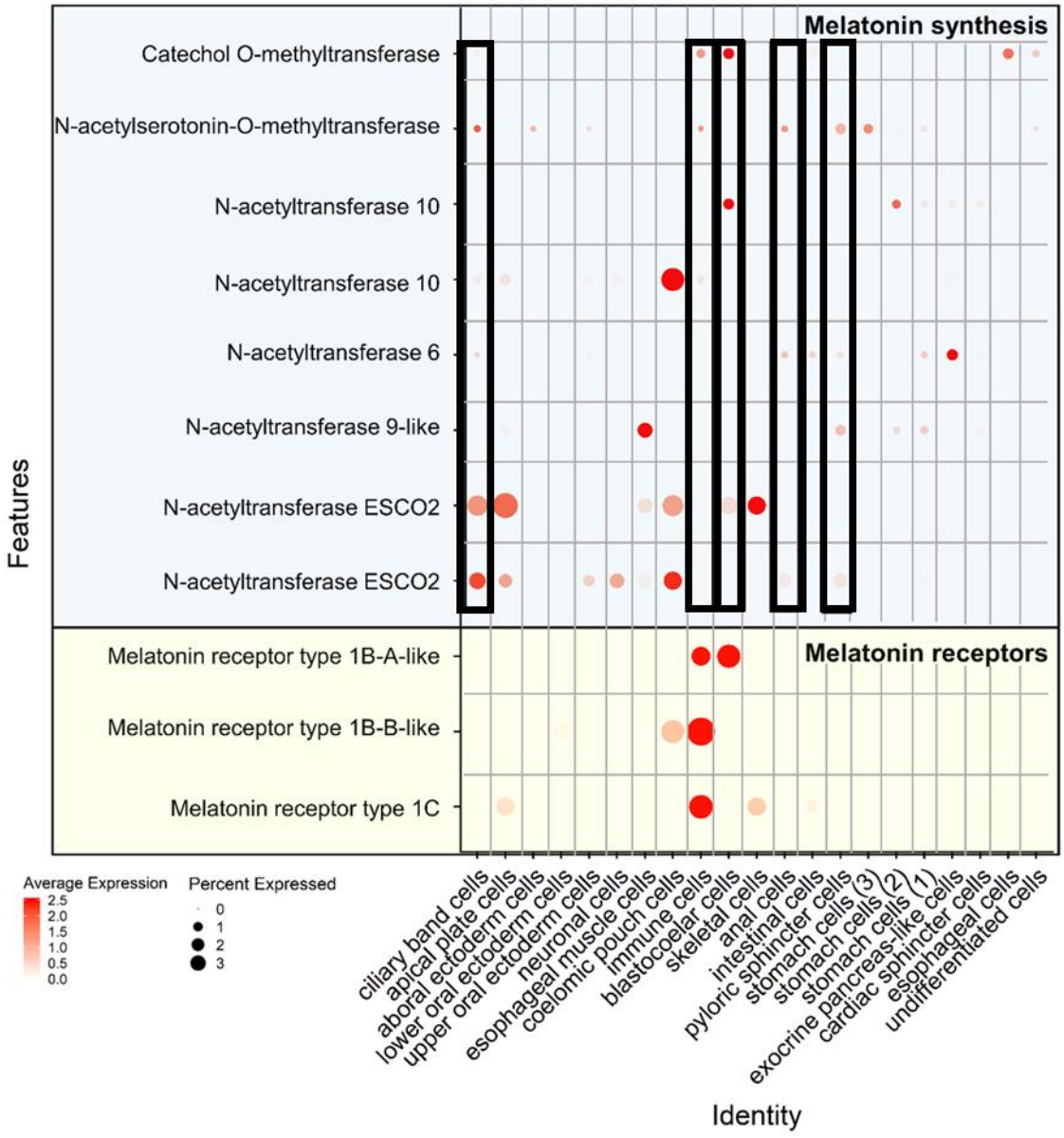


Figure 4.14. Melatonin biosynthetic and signaling pathway genes at 3 dpf larva. Dotplot showing the average expression of genes involved in the biosynthesis and signaling pathway of melatonin. Bars indicate the putative broad cell types, in which transcripts for the necessary biosynthetic enzymes are co-localized and thus can putatively produce melatonin.

4.2.2. Dopamine signaling

Dopamine is a neuromodulator produced by several various cell populations of the vertebrate nervous system, and it has been reporting that the role of dopamine in chemical signaling predates the divergence of deuterostomes (Yamamoto and

Vernier 2011, Baudonnat et al. 2013)). In mammals, dopamine has been found to have a vital role in sensorimotor programming, learning, memory, and many cognitive and autonomous functions (Baudonnat et al., 2013).

In sea urchins, dopamine is produced throughout most of the embryonic and larval development, and its role has been disentangled in great detail. Pharmacological inhibition of dopamine decarboxylase activity in embryos and larvae of *H. pulcherrimus* resulted in embryos and larvae with reduced swimming ability (Katow et al. 2010), while in *S. purpuratus,* dopaminergic signaling has been implicated with the response of larvae to food availability (Adams et al. 2011). When food is absent, they develop longer arms than otherwise. Adams et al. (2011) also demonstrated that the dopamine signaling pathway components' spatial and temporal expression is in line with the feeding behavior of the larva and that algae-induced dopamine signaling inhibits arm elongation. Moreover, Burke and colleagues found that externally provided dopamine-induced the sand dollar *Dendraster excentricus* larvae's metamorphic process, suggesting that the dopaminergic system is also implicated in echinoderm metamorphosis (Burke, 1983).

Plotting of the genes involved in dopamine signaling verified that tyrosine hydroxylase and dopamine decarboxylase, both involved in dopamine biosynthesis, are differentially expressed in neuronal cells (Fig. 4.15). According to the same scRNA-seq analysis, it seems that dopamine receptors are highly distributed in ectodermally derived cell types such as the ciliary band and apical plate cells and differentially expressed in mesodermal cell types such as skeletal, immune, and blastocoelar cells. The presence of dopamine receptors in the ciliary band and apical plate cell types, which bare a significant number of motile cilia, is in agreement with the hypothesized role of dopamine in regulating sea urchin swimming behavior. On a similar note, the presence of transcripts of dopamine receptors in the skeletal putative broad cell types is an additional validation of the Adams et al. study as well as indicative of a possible role of endogenous dopamine in regulating skeletal growth.

In the same analysis, dopamine β hydroxylase, which is involved in the formation of noradrenaline from dopamine, is found to be enriched in the apical plate, neuronal, and blastocoelar putative broad cell types while the same cell types bear an adrenergic receptor. Until now, there was no information on the expression of *dopamine β hydroxylase*, and thus this study offers the starting point for future

studies aiming to disentangle the role and evolution of noradrenalin signaling in sea urchin.

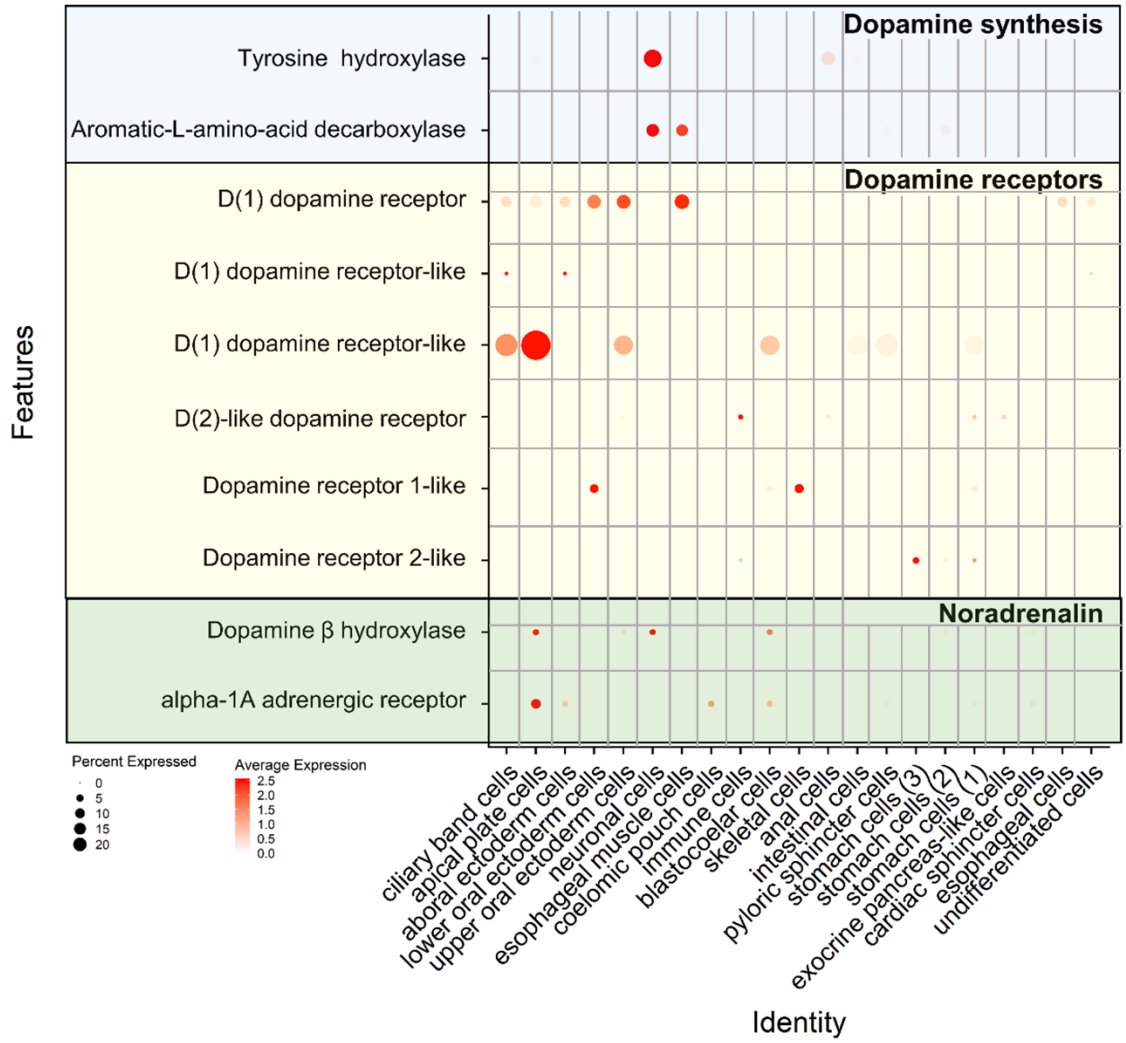


Figure 4.15. Dopamine and noradrenalin signalings at 3 dpf larva. Dotplot showing the average expression of genes involved in the biosynthesis, inactivation, and signaling cascade of dopamine and noradrenaline.

4.2.3. GABA signaling

GABA, or else known as γ-aminobutyric acid, is a well-known neurotransmitter with an inhibitory role in the central nervous system of mammals (Tano et al., 2009). The biosynthesis of GABA depends on the enzyme *glutamic acid decarboxylase (GAD)* that converts glutamate to GABA. GABA has been found to function through post-synaptic binding to two classes of receptors GABAA and GABAB. (Simeone et al., 2003 (Wang et al., 1999). The only reports of the GABAergic signaling system in sea urchin come from studies carried out in the sea urchin *H. pulcherrimus.* These

studies demonstrated that GABA is produced by various sea urchin cell types during embryonic development and mostly by blastocoelar cells. Similarly, *Hp-Gad,* both at transcript and protein levels, was detected in endodermal cell types and blastocoelar cells. Furthermore, pharmacological inhibition of *GAD* resulted in larvae with reduced swimming ability and sea urchin juveniles with lower motility (Katow et al. 2013, Katow et al. 2020).

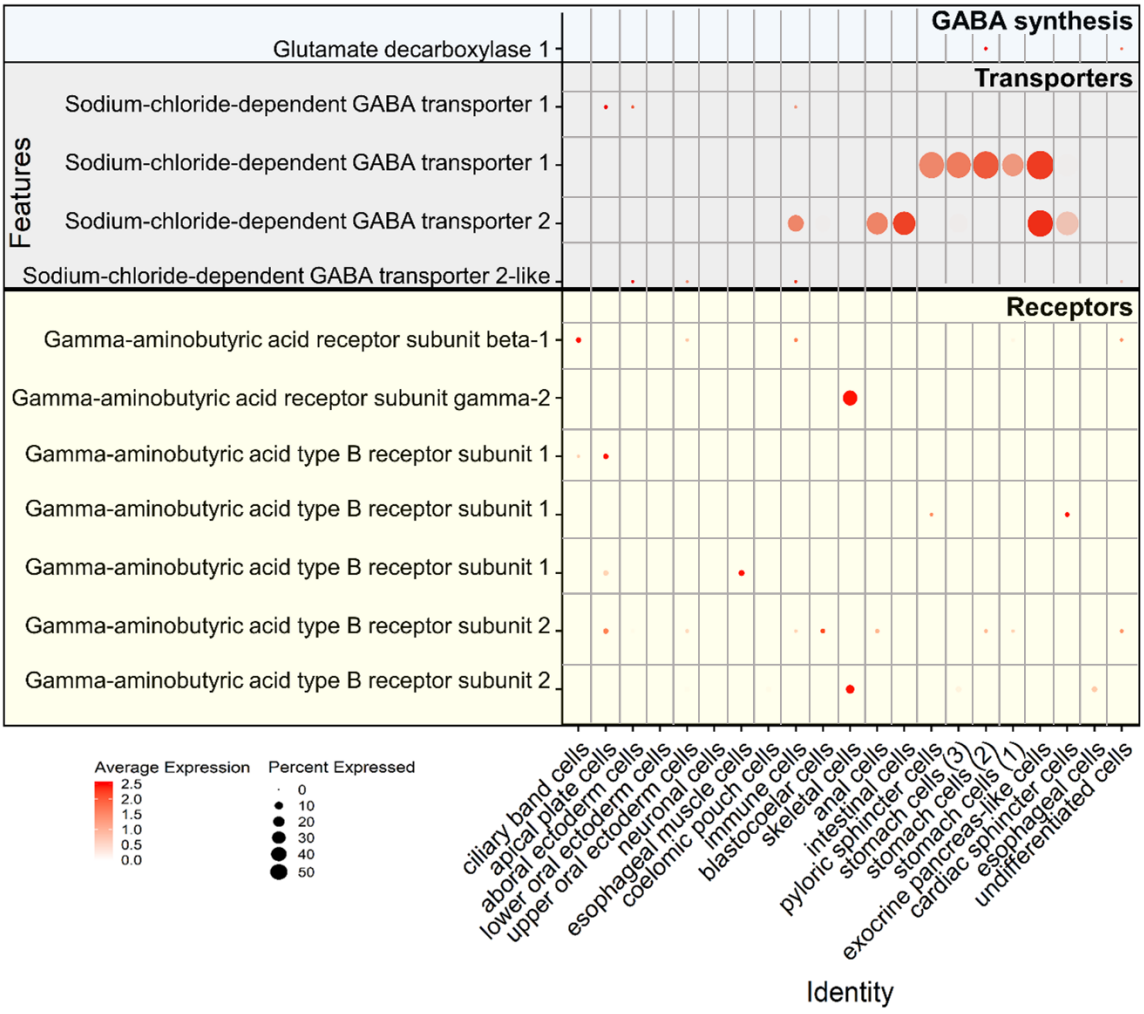


Figure 4.16. GABA signaling at 3 dpf larva. Dotplot showing the average expression of genes involved in the biosynthesis, transport and signaling cascade of GABA.

Plotting of the *S. purpuratus* homolog *Sp-Gad* and the GABA signaling components revealed very low expression of *Sp-Gad* that was restricted in one of the stomach putative cell types. At the same time, the GABA-associated receptors seem to be also transcribed at low levels, mostly in the skeletal, apical plate, ciliary band, and esophageal muscle cell types (Fig. 4.16). Interestingly sodium and chloride-dependent GABA transporters were found to be also differentially expressed in

almost all endodermally derived cell types as well as the immune cells (Fig. 4.16). *Sp-Gad*'s presence in the stomach cell type agrees with one of the expression domains described in *H. pulcherrimus* (Katow et al. 2014), although in *S. purpuratus,* no expression was found in blastocoelar cells, the primary domain of *H. pulcherrimus,* where it has been hypothesized that GABA functions (Katow et al. 2013). On the other hand, the expression of the receptors in the ciliary band and apical plate cell types hints at a possible role in ciliary locomotion control in *S. purpuratus* as in *H. pulcherrimus*.

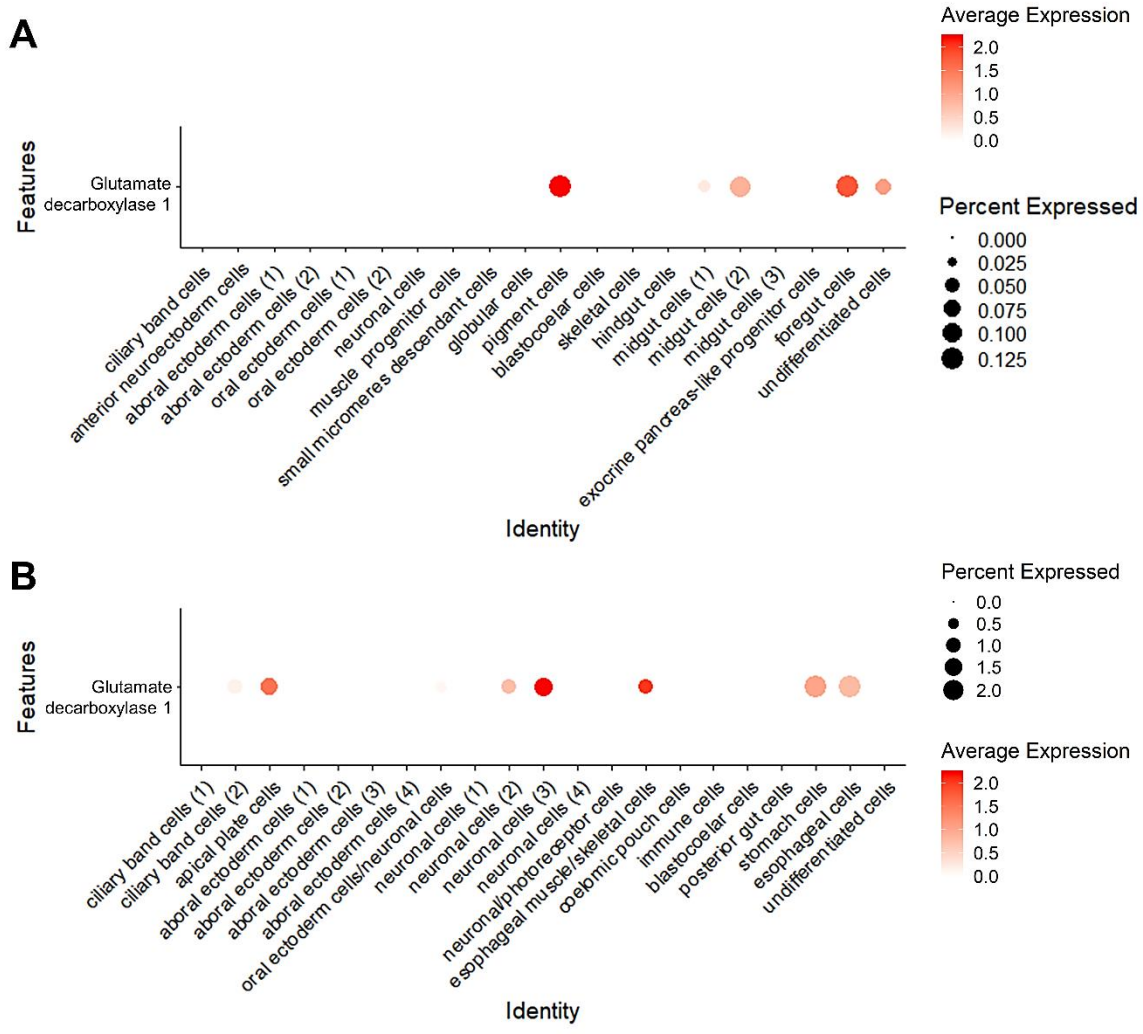


Figure 4.17. Glutamate decarboxylase expression across development. Dotplot showing the average expression of *glutamate decarboxylase* at 2 dpf (A) and 5 dpf (B).

To understand whether the absence of Sp-Gad transcripts in blastocoelar cells is a cross-species difference or an artifact of the scRNA-seq data at this time-point (3 dpf), *Sp-Gad* was also plotted in the other two datasets. In the case of the 2 dpf gastrula dataset, *Sp-Gad* was expressed at even lower levels and enriched in three

endodermally derived cell types (midgut 1, midgut 2, and foregut) as well as in the pigment cell cluster (Fig. 4.17 A). Similarly, at 5 dpf, *Sp-Gad* transcripts were found to be differentially expressed in the stomach and esophageal cells and in additional domains such as neuronal, skeletal, and ciliary band cells (Fig. 4.17 B). In all cases, no blastocoelar cell expression could be found. According to this transcriptomic-based analysis, *Sp-Gad's* expression in the endoderm seems conserved in *S. purpuratus* and *H. pulcherrimus*, although the blastocoelar cell expression in *H. pulcherrimus* might be unique of this species. Further studies are needed to *in vivo* characterize *Sp-Gad's* expression pattern and validate the single-cell predictions before reaching a definite conclusion on the evolution of this gene and its signaling.

4.2.4. Acetylcholine signaling

Acetylcholine (Ach) derives from acetyl coenzyme A and choline, and the enzyme choline acetyltransferase (Chat) catalyzes this reaction. Acetylcholine has been found to function in the peripheral nervous system of mammals, and mostly it has been detected at synapsis in the ganglia visceral motor system (Zaninetti et al. 1999, Nishimaru et al. 2005). In several invertebrates, such as mollusks, annelids and hemichordates, acetylcholine has been implicated in the control of ciliary beating d (Marinkovic et al. 2020), and in echinoderms, it has been proposed that acetylcholine has a role in decreasing the ciliary beating frequency of sea urchin larvae and causing muscle contraction in sea urchin tube feet (Florey et al. 1975, Soliman 1984).

According to the scRNA-seq analysis, *Sp-Chat* transcripts are found to be differentially expressed in the neuronal putative cell type and the putative ACh receptors seem to be expressed by several ectodermally and mesodermally derived cell types and especially in the neurons. The broad distribution of the receptors suggests that ACh signaling might have a broad spectrum of functions in regulating diverse processes (Fig. 4.18). Strikingly, what is evident from this analysis is that the majority of the enzymes that inactivate acetylcholine, acetylcholinesterases, are enriched in endodermally derived cell types (Fig. 4.18), suggesting that acetylcholine degradation takes place in the digestive tract. This would suggest that the gut of the sea urchin larva is responsible for reducing the available amounts of acetylcholine and thus participating in the control and duration of ACh signaling.

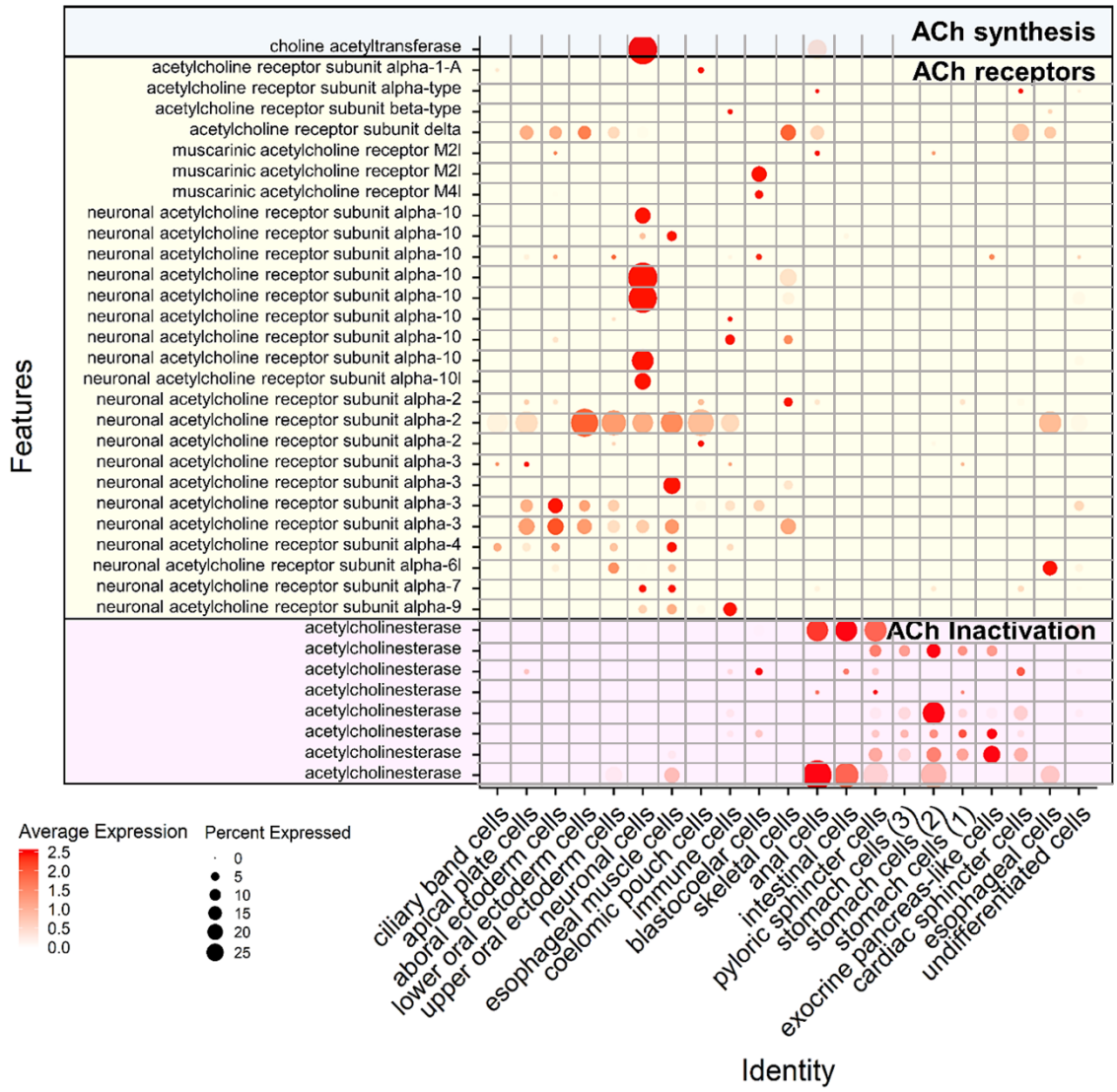


Figure 4.18. ACh signaling at 3 dpf larva. Dotplot showing the average expression of genes involved in the biosynthesis, signaling cascade, and inactivation of ACh.

4.2.5. Endocannabinoid signaling

As previously mentioned, the endocannabinoid signaling (ES), alongside nitric oxide, belongs to the unconventional neurotransmitter group. ES is based on unsaturated fatty acids diffusing through a cell's membrane and activating a signaling cascade on nearby cells that express cannabinoid receptors. Anandamide (AEA) and 2-arachidonylglycerol (2-AG) have been identified as endocannabinoids, and their biosynthetic pathway has been described in great detail (Fig. 4.19). Endocannabinoids are found to function through binding to several receptor types, including the traditional cannabinoid CB1 and CB2 receptors. Additionally, endocannabinoids can interact and initiate their signaling through the *peroxisome proliferator-activated receptors (Ppars)* (O'Sullivan 2007, O'Sullivan 2016). Once

endocannabinoids exhibit their function, they are hydrolyzed by fatty acid hydrolases.

ES is involved in the regulation of many biological processes in vertebrates spanning from mediating synaptic plasticity at both excitatory and inhibitory synapses (Castillo et al. 2012) to ensuring proper vision (Martella et al. 2016) and participating in the central control of reproduction (Cottone et al. 2013). Historically, it had been hypothesized that endocannabinoid signaling was a vertebrate innovation, although several studies that took place over the last two decades demonstrated that the signaling is also active in invertebrate deuterostomes (Salzet and Stefano 2002, Elphick et al. 2003, Elphick 2012, Lemak 2012). The endocannabinoid system in invertebrates has multiply roles spanning from control of reproduction, feeding behavior and neurotransmission to immune defense (Salzet and Stefano 2002, Lemak 2012).

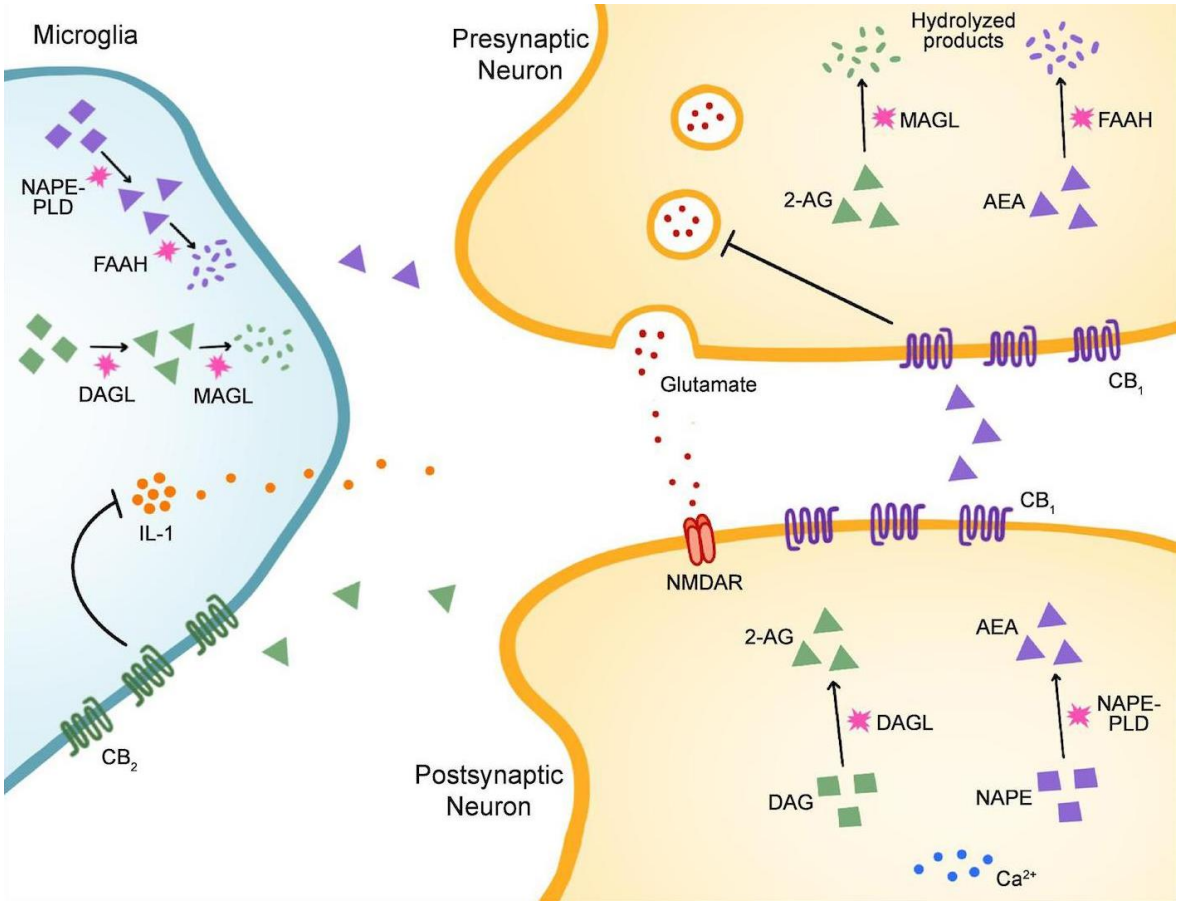


Figure 4.19. Endocannabinoid signaling. Schematic representation of the biosynthesis, signaling cascade, and hydrolysis of A-AG and AEA [adapted from (Araujo et al. 2019)].

The only evidence of ES functioning in sea urchin comes from studies in which ES has been implicated with the control of the fertilization process and polyspermy prevention (Chang et al. 1993; Schuel et al. 1994). Plotting for the homologs of the vertebrate genes encoding for components of ES revealed that all the necessary machinery for ES is present at 3 dpf *S. purpuratus* larva (Fig. 4.20). In more detail, transcripts for three *diacylglycerol lipases* involved in the biosynthesis of 2-AG are found in two oral ectoderm, neuronal, coelomic pouch, blastocoelar, pyloric sphincter, and exocrine pancreas like cells. Transcripts for *N-acyl phosphatidylethanolamine phospholipase D* responsible for the production of AEA are localized in the apical plate, neuronal, and esophageal muscle cells. Taken together, these data suggest that the sea urchin larva can produce the 2-AE and AEA endocannabinoids from both neuronal and non-neuronal cell types, hinting a broad spectrum of ES function in echinoderms. The only receptors found in the scRNA-seq datasets are non-traditional nuclear hormone peroxisome proliferator-activated receptors (*Pparα and Pparγ*). While the *S. purpuratus* genome encodes for a gene for the cannabinoid receptor 1, no transcripts could be detected in the scRNA-seq datasets. However, the spatial distribution of Ppars in skeletal and several endodermally derived cell types (Fig. 4.20) suggests that the sea urchin endocannabinoid signaling could be the mediator of a cross-talk between several cell types, such as neurons, with the digestive tract. In support of this, the degrading enzyme for 2-AG *monoglyceride lipase,* as well as the enzymes *fatty acid amide hydrolases* involved in the hydrolysis of both 2-AG and AEA, are enriched in endodermally derived cell types suggesting that the digestive tract is strongly receiving the endocannabinoids and it is the prominent place of their hydrolysis.

Moreover, based on endocannabinoids' ability to diffuse for short distances, the ES may also be used as means of communication between different cell types patterning the sea urchin gut. For instance, *diacylglycerol lipase β like* is expressed in the pyloric sphincter and exocrine pancreas-like cell types suggesting that 2-AG is produced there as well. The distribution of *Ppars* transcripts in intestinal and stomach cells that are in close proximity to the two aforementioned cell types respectively implies a possible directional communication initiated by the pyloric sphincter cells signaling to the intestinal cells and by the exocrine pancreas-like cells signaling to the stomach cells. Future studies will verify the *in vivo* expression patterns of the ES genes and the function of ES in sea urchin.

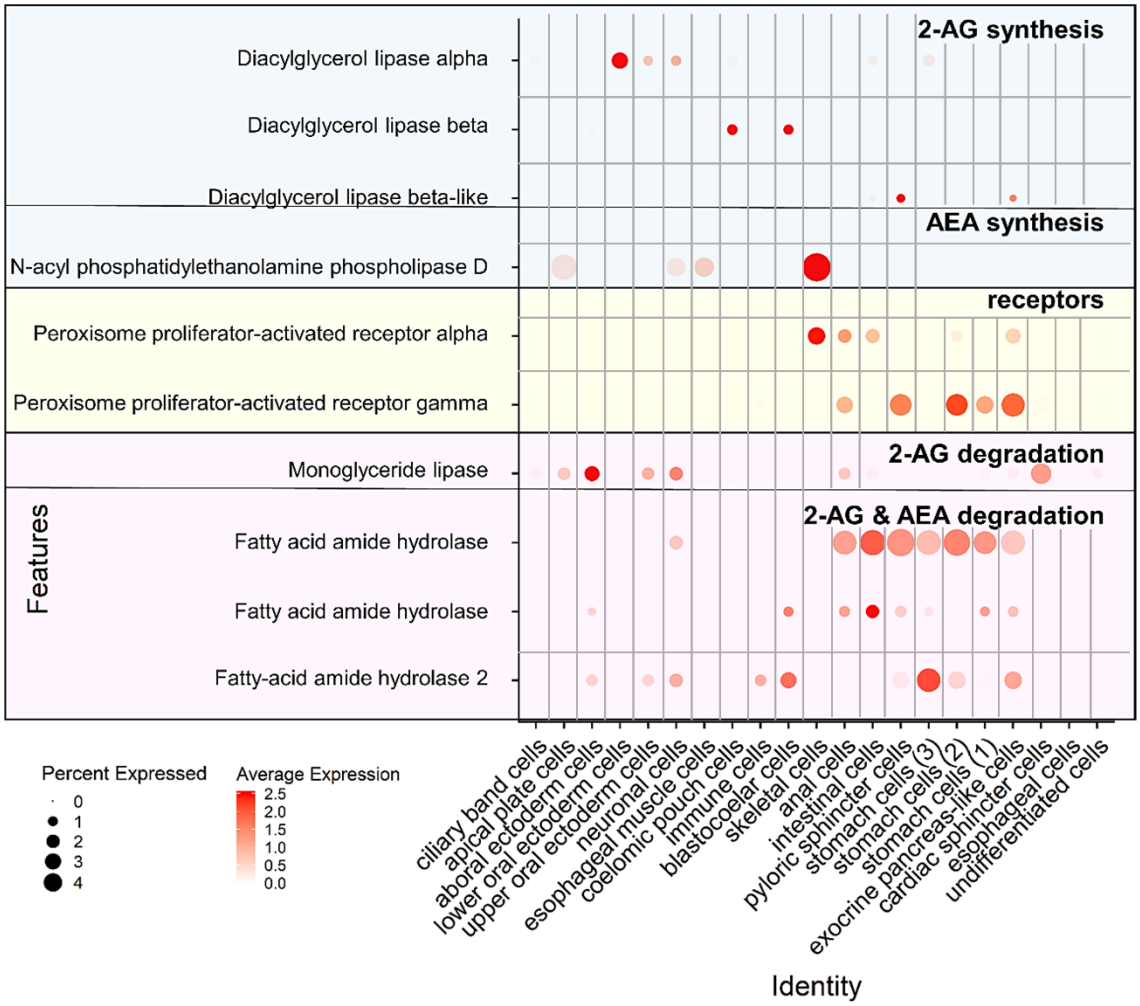


Figure 4.20. Endocannabinoid system components in sea urchin larva. Dotplot showing the average expression of genes involved in the biosynthesis, signaling cascade, and degradation of cannabinoids.

4.2.6. Nitric oxide

One of the most well-known unconventional neurotransmitters is nitric oxide (NO). NO is a gas produced by the enzyme nitric oxide synthase (Nos) that converts arginine into citrulline while generating NO. The generated NO can diffuse through the plasma membrane of the cell producing it and acts on neighboring cells, which is a feature allowing NO to regulate multiple processes in a great spectrum of nearby cell types. Due to its ability to penetrate cells without any selectivity, e.g., binding to a receptor, NO is often considered a second messenger rather than a neuromodulator.

Once in the target cell, NO can function either through the canonical or the non-canonical pathway (Fig. 4.21).

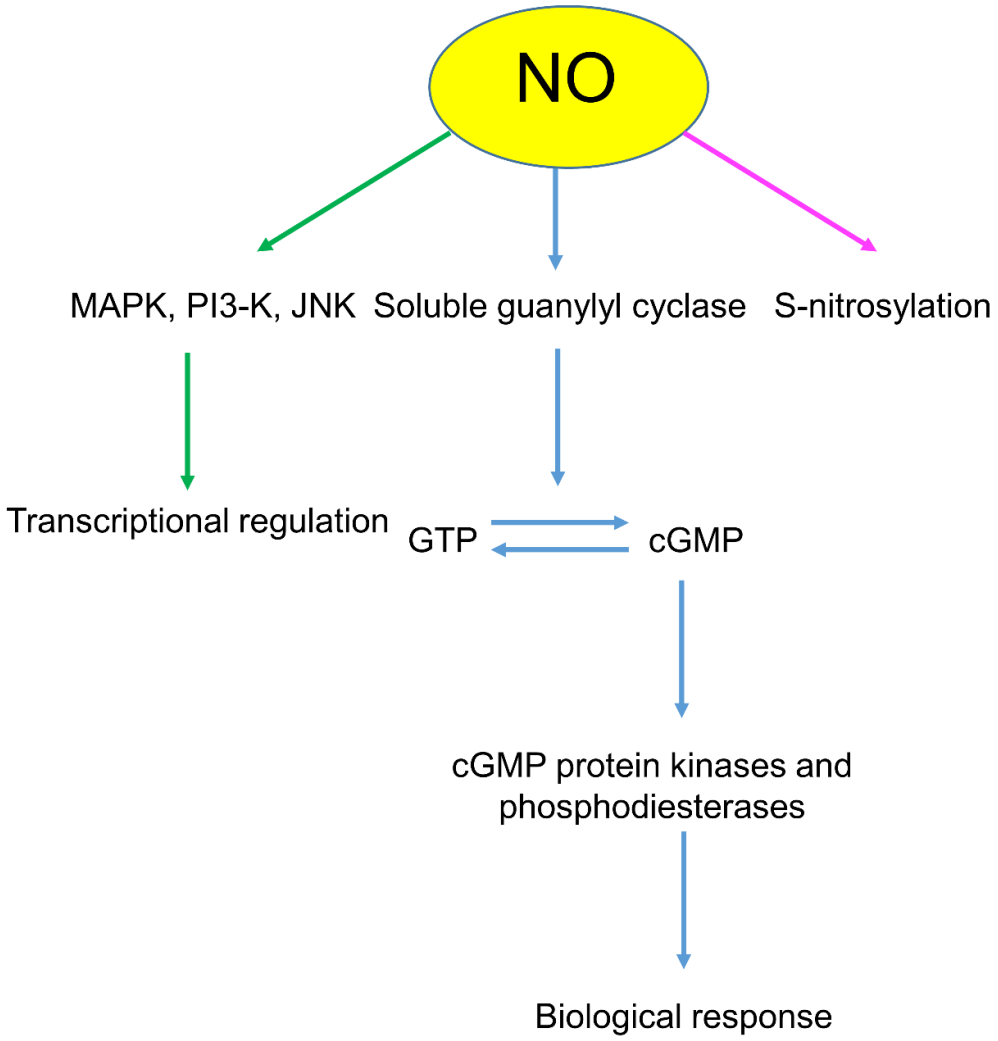


Figure 4.21. Summary of the NO signaling pathways. Overview of the different ways by which NO exhibits its function.

Two pathways have been associated with NO canonical signaling. The first one includes the interaction of NO with kinases such as mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3-K), and the protein kinase JNK that results in transcriptional regulation (Park et al. 2000, Pei et al. 2008, Doronzo et al. 2011, Castellano et al. 2014). The second pathway includes activating the enzyme guanylyl cyclase by NO, which then produces the second messenger cyclic guanosine monophosphate (cGMP). Next, the cGMP binds and activates its target, which is a cGMP-dependent protein kinase. As a result, the activated kinase phosphorylates target proteins and initiates a signaling cascade (Yoshioka et al. 2006, Martinez-Ruiz et al. 2011, Wang et al. 2017). Apart from the kinase, cGMP can bind to ligand-gated ion channels influencing neuronal signaling. Apart from initiating signaling pathways resulting in gene expression regulation at a

transcriptional level, NO can function post-translationally the S-nitrosylation. S-nitrosylation is a protein modification that results from NO directly modifying sulfhydryl protein residues, and there have been reports that it is a widespread modification comparable to phosphorylation (Stamler et al. 1992, Broillet 1999).

Due to the ability of NO to diffuse in multiple cell types and affect them in many different ways, it comes as no surprise that according to literature, NO is involved in the regulation of many processes ranging from morphogenesis, homeostasis, physiology, immune system defense, neuronal communication to specialized organ function in both vertebrates and invertebrates (Martinez 1995, Tatsumi et al. 2004, Palumbo 2005, Rivero 2006, Cristino et al. 2008, Kolluru et al. 2008, Galdino et al. 2015, Annona et al. 2017).

In echinoderms, NO has been found to function as a general myorelaxant. In the starfish *Asterias rubens* NO induces the relaxation of the cardiac stomach (Elphick and Melarange 1998), while in sea urchin larva of *H. pulcherrimus* endodermally-derived NOS-positive neurons produce NO that regulates the relaxation of the pyloric sphincter (Yaguchi and Yaguchi 2019). The *S. purpuratus* genome contains two Nos genes, and phylogenetic analysis across metazoans revealed that these two genes correspond to *NosA* and *NosB* and that NOSA contained a PDZ domain, indicative of a neuronal function (Andreakis et al. 2011).

Based on the putative broad spectrum of function in sea urchins, and since the only report of nitric oxide function concerns the *NosA* gene (Yaguchi and Yaguchi 2019), I decided to further characterize the nitric oxide pathway function during *S. purpuratus* development. Plotting the average expression of *NosA* and *NosB* in all three time-points showed that *NosA* transcripts are present in all of them and that it is differentially expressed in endodermally derived cell types (Fig. 4.22) suggesting a role in the regulation of the gut function similar to what has been demonstrated in *H. pulcherrimus* (Yaguchi and Yaguchi 2019). In particular, at gastrula stage, *NosA* transcripts are detected apart from the endoderm in various ectodermally derived cell types that correspond to the anterior neuroectoderm, oral and aboral ectoderm. NosB transcripts are detected only at the pluteus stages, which is in agreement with data coming from bulk RNA-seq (Fig. 4.22). NosB transcripts are localized mostly in ectodermally derived cell types of pluteus stages, including ciliary band apical plate and neurons, suggesting a neuronal role.

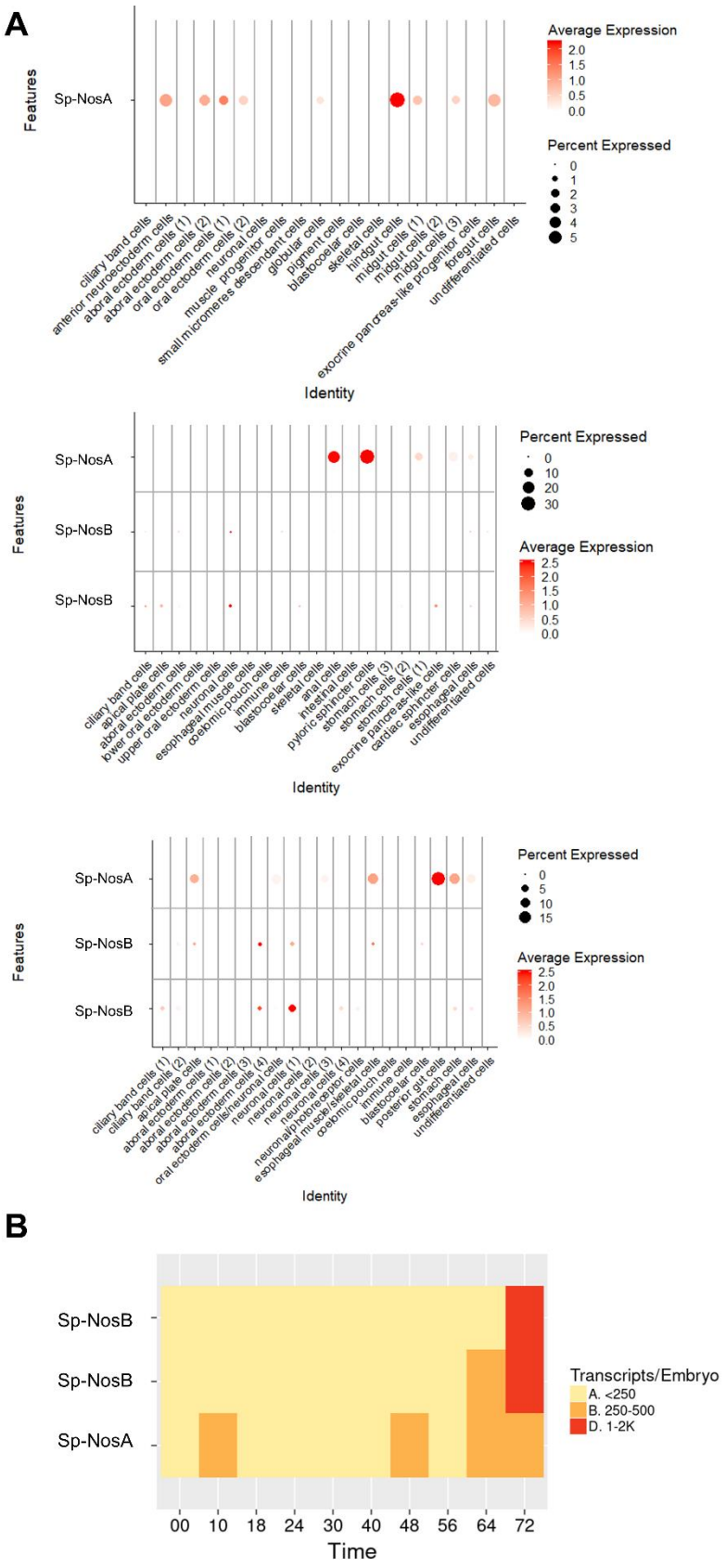


Figure 4.22. Distribution of Nos transcripts during development. A) Dotplots showing the average expression of Sp-NosA and Sp-NosB genes at 2, 3, and 5 dpf developmental stages. B) Temporal expression of Sp-NosA and Sp-NosB genes during early embryonic and larval development as calculated from transcriptomic data (Tu et al. 2014).

To validate the scRNA-seq predictions, I cloned the *Sp-NosA* and *Sp-NosB* genes with the help of Dr. Caccavale and performed FISH and IHC experiments using a universal Nos antibody (uNos) on several developmental time-points. At 2 dpf gastrula, *Sp-NosA* transcripts were enriched in the digestive tract, especially in two bilateral pairs of cells located at the lowest level of the foregut (Fig. 4.23 A1). Moreover, a few NosA positive cells were detected in the anterior neuroectoderm and aboral ectoderm (Fig. 4.23 A1). At 3 dpf pluteus stage, *Sp-NosA* signal was detected in endodermal cells close to the anus, pyloric sphincter, parts of the stomach, and a ring of cells at the lowest level of the esophagus (Fig. 4.23 A2, A3) that resembles the expression of *Sp-NosA* in the bilateral patch at gastrula stage. FISH for *Sp-NosB* failed to detect any transcripts (data not shown) at the 2 dpf gastrula stage validating the scRNA-seq and transcriptomic predictions, while at 3 dpf pluteus stage, *Sp-NosB* was detected in two cells distally to the apical organ as well as in scattered ciliary band and apical plate cells (Fig. 4.23 A4). IHC experiments using the universal Nos antibody that recognizes the conserved domains of different Nos isoforms recapitulate the same expression patterns identified by FISH. At gastrula stage, uNos positive cells are detected in the anterior neuroectoderm and ciliary band and aboral ectoderm (Fig. 4.23 A5), while at 3 dpf pluteus uNos positive cells are found in all cell types expressing *Sp-NosA* and *Sp-NosB,* showing that indeed the antibody can recognize both Nos proteins (Fig. 4.23 A6). Similarly, at 5 and 7 dpf pluteus, uNos signal was found in the digestive tract, primarily in the posterior domain, including the pyloric sphincter, intestine, and cells around the anus as well as cells distally to the apical organ (Fig. 4.23 A7, A8). Co-staining of 1-week post fertilization larvae with uNos and the neuronal marker SynB revealed that those cells in close proximity to the apical organ are neurons (Fig. 4.23 A8).

Since FISH and IHC validated the scRNA-seq predictions regarding the spatial expression of the two Nos genes, I set out to investigate in which cell types NO is localized. To achieve this goal, I used 4-amino-5-methylamino-2',7'-difluorofluorescein diacetate (DAF-FM-DA), which is a cell-permeable and non-fluorescent reagent that once it reacts with NO forms the fluorescent compound benzotriazole. This compound has been widely used to detect NO in whole organisms, tissues, and cell cultures (Rhinehart and Pallone 2001, Kashiwagi et al. 2002, Pandey et al. 2015, St Laurent et al. 2015, Annona et al. 2017). DAF-NO mediated fluorescence was detected in the ciliary band, apical plate, and the whole

digestive tract of pluteus larvae at 3 and 4 dpf (Fig. 4.23 A9, A10, A11, A12). Interestingly, DAF-related fluorescence was never detected in any mesodermal cell types within the blastocoel or the aboral ectodermal epithelium, indicating that NO signaling is not taking place there.

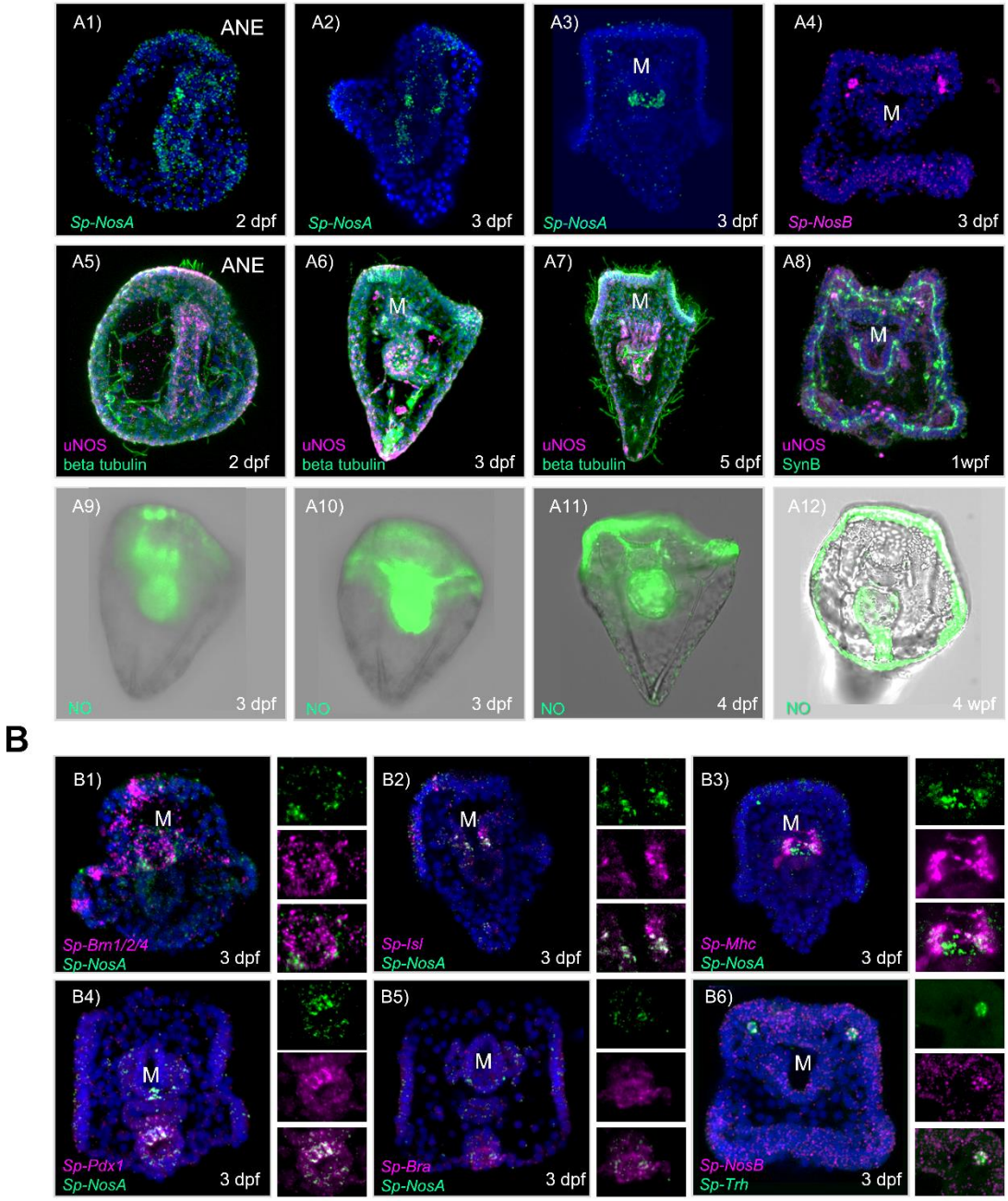


Figure 4.23. *In vivo* spatial localization of Nos genes and NO. A) Spatial localization of Sp-NosA and Sp-NosB transcripts, Nos protein, and NO. FISH using specific antisense RNA probes for Sp-NosA (A1-A3) and NosB (A4). IHC detection of Nos using a universal Nos antibody at 2 dpf (A5), 3 dpf (A6), 5 dpf (A7), and 1 wpf (A8) developmental stages. NO visualization by DAF at 3 dpf (A9-10) and 4 dpf (A11-12) larvae. B) Molecular characterization of the Sp-NosA and Sp-NosB positive cells. Double FISH for Sp-NosA and Sp-Brn1/2/4 (B1), Sp-Isl (B2), Sp-Mhc (B3), Sp-Pdx1 (B4), Sp-Bra (B5) as well as for Sp-

NosB with Sp-Trh (B6). Nuclei are labeled with DAPI (in blue). Confocal images are stacks of merged Z sections. ANE, anterior neuroectoderm; M, mouth.

Moreover, the DAF fluorescence enrichment and thus NO localization in ectodermal and endodermal cell types is in line with the complementary broad expression domains of the NosA and NosB in endoderm and ectoderm, respectively. Unfortunately, based on the ability of NO to diffuse in a distance of 30-100 μm combined with the average size of the larva across the anterior to the posterior axis that is approximately 120 μm, the exact tracing of NO source per cell type cannot be concluded.

Next, to further characterize the Nos producing cells, I performed double FISH with additional molecular markers (Fig. 4.23 B). This set of experiments revealed that NosA is co-expressed with the neurogenic marker *Sp-Brn1/2/4* and the neurogenic marker revealed by this thesis *Sp-Isl* in the lower esophageal domain (Fig. 4.23 B1, B2). These putative NosA positive neurons are in close proximity to the esophageal muscle apparatus, as revealed by the co-staining using a specific antisense probe for *Sp-Mhc* (*Myosin heavy chain*) (Fig. 4.23 B3). Since Sp-NosA is expressed in several endodermal domains, the markers *Sp-Pdx1* and *Sp-Bra* for the pyloric sphincter and cells around the anus, respectively, were used. Double FISH revealed that *Sp-NosA* is co-expressed with those two transcription factors in the domains mentioned above, validating the single-cell predictions (Fig. 4.23 B4, B5). Based on the Sp-NosB positive cells' peculiar position flanking the apical organ that resembles neurons that produce the neuropeptide Trh and the scRNA-seq predictions, double FISH of *Sp-NosB* and *Sp-Trh* were performed, that validated the identity of those cells as the Trh neurons (Fig. 4.23 B6).

After establishing the molecular signature and spatial localization of the Nos expressing genes and NO positive cells, functional studies were carried out to decipher the role of NO signaling in the sea urchin *S. purpuratus.* To do so, two different approaches were followed. In order to inhibit NO synthesis, 1-(α,α,α-trifluoro-o-tolyl)-Imidazole (TRIM) was used, as it has been demonstrated that it successfully inhibits proper NOS (Nitric oxide synthase) activity in a spectrum of organisms (Fukunaga et al. 2000, Haga et al. 2003, Annona et al. 2017). TRIM treatments were carried out at 2 dpf after the gastrulation process was completed, and the embryos were let to develop until the pluteus stage. To knock down the individual *Sp-NosA* and *Sp-NosB* gene expression, specific morpholino antisense

oligonucleotides were injected into *S. purpuratus* fertilized eggs and were let to develop until the pluteus stage.

Inhibiting Nos activity or knocking down the expression of either of Nos genes resulted in distinct, although similar phenotypes. The efficiency of the TRIM treatments and the morpholino injections was assessed by using DAF to visualize the amount of NO produced and immunohistochemical detection of the Nos protein, respectively (Fig.4.24). At 3 dpf TRIM treated larvae (Fig.4.24 A2, A7, A12) were smaller in size in respect to the controls (Fig.4.24 A1, A6, A11), the overall shape of the larva is altered, and very little DAF-NO related fluorescence was detected (Fig. 4.24 A12). Morphological deficits on the ciliary band and post-oral arm formation and aboral ectoderm size are evident with confocal microscopy using the marker b-tubulin (Fig. 4.24 A7) that allows the visualization of multiple structures as well as light microscopy. Knockdown (KD) of Sp-NosA resulted to a comparable phenotype in regards to the one obtained by Nos inhibition and immunostaining for the Nos protein did not show any signal, implying that the Sp-NosA knockdown was successful (Fig.4.24 A3, A8, A13). The Sp-NosA KD larvae appear to be smaller in size, although partial formation of post-oral arms is evident. On the other hand, Sp-NosB KD dramatically affected embryonic development since the Sp-NosB morphants show signs of embryonic arrest, and immunohistochemical detection of Nos was restricted to endodermal structures implying that it detects only NosA (Fig.4.24 A4, A9, A14). The overall shape of the larva is gastrula-like, while skeletal structures are disorganized. An interesting aspect of the Sp-NosB KD phenotype is that the skeletal rods show signs of bifurcation, which is a feature of the larval skeleton but completely disorganized in terms of spatial distribution. This is probably the effect of the ectodermal defects that do not provide PMCs with the proper cues to migrate. The digestive tract of the Sp-NosB morphants seems like a straight tube with no obvious compartmentalization. Co-injection of the two morpholinos targeting Sp-NosA and Sp-NosB resulted in larvae phenotypically similar to the TRIM treated and Sp-NosA morphant ones (Fig.4.24 A5, A10, A15), suggesting that the extreme phenotype of the Sp-NosB KD alone is probably a result of technical issues that occurred during the microinjection procedure (morpholino concentration, the microinjection procedure itself). The inhibition of Nos after TRIM treatment was carried out in five biological replicates, and a reproducible phenotype was obtained in every set of experiments allowing me to gain confidence that the embryonic deficits are a direct result of Nos inhibition. On the contrary, the morpholino injection

experiments for *Sp-NosA, Sp-NosB*, and *Sp-NosA* and *Sp-NosA/Sp-NosB* were carried out only in larvae from one batch consisting of one biological replicate, and thus no definite conclusion can be drawn. Nonetheless, most of these preliminary gene perturbation results are consistent with the TRIM generated phenotype highlighting the involvement of NO in morphogenesis.

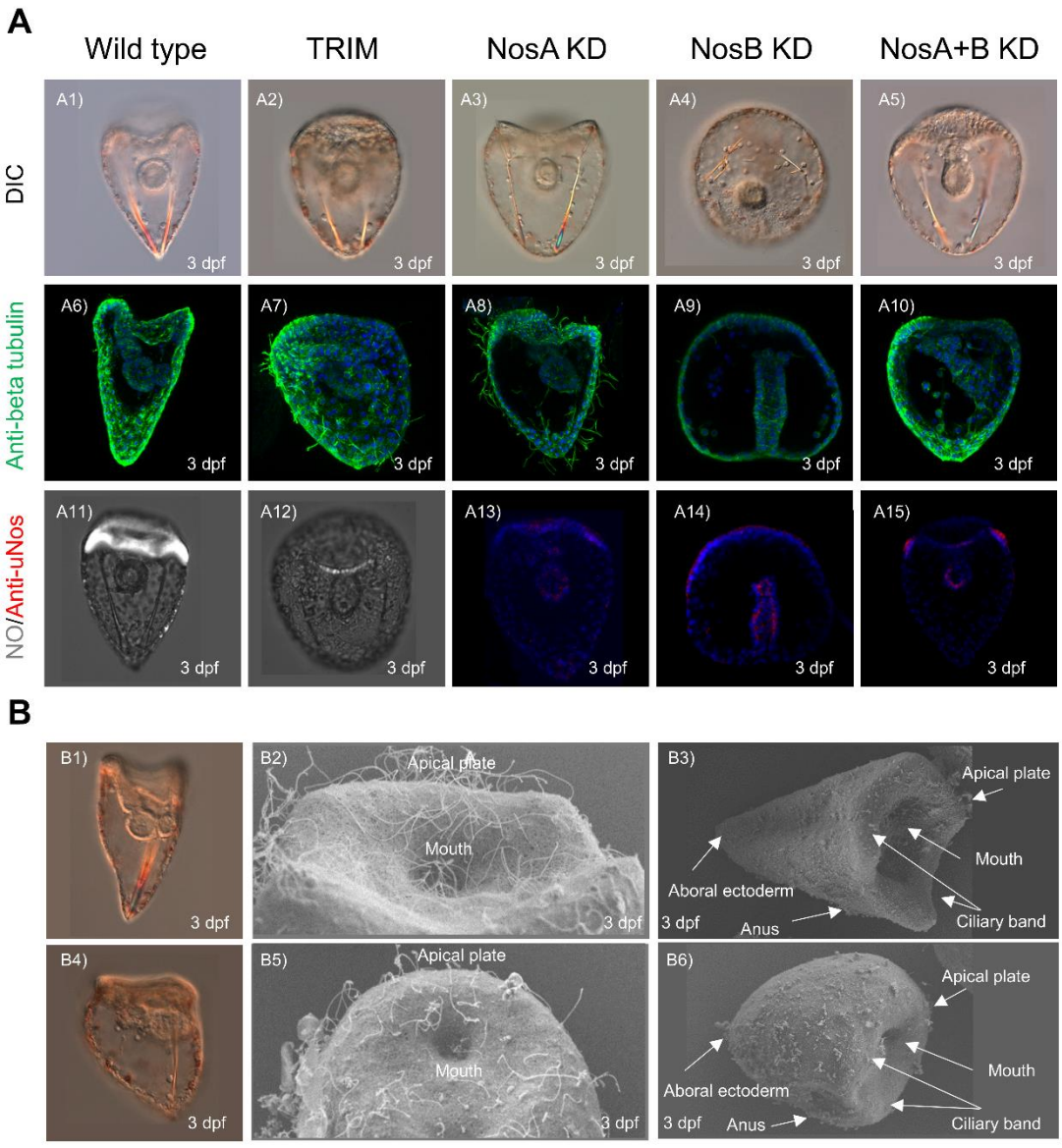


Figure 4.24. Inhibition of NO production. A) Morphological and molecular characterization of NO deficient of 3 dpf larvae. DIC images of wild type (A1), TRIM treated (A2), NosA KD (A3), NosB (A4), and NosA/NosB KD (A5) larvae. Immunohistochemical detection of beta-tubulin in wild type (A6), TRIM treated (A7), NosA KD (A8), NosB (A9), and NosA/NosB (A10) KD larvae. NO localization based on DAF-NO mediated fluorescence in wild type (A11) and TRIM treated (A12) larvae. Immunostaining using the uNOs antibody on NosA KD (A13), NosB (A14), and NosA/NosB (A15) KD larvae. B) Morphological characterization of TRIM treated 3 dpf larvae. DIC images of wild type (B1) and TRIM treated (B4) larvae. SEM images of wild type (B2, B3) and TRIM (B5, B6) treated larvae. Nuclei are labeled with DAPI (in blue). Confocal images are stacks of merged Z sections.

Based on the high confidence obtained regarding the TRIM treated larvae phenotype and its reproducibility, I set out to further characterize the morphological and molecular consequences of the NO production inhibition. SEM imaging of 3 dpf TRIM treated larvae, and controls resulted in similar conclusions as the ones obtained from light microscopy observations (Fig. 4.24 B). Control larvae (Fig. 4.24 B1, B3, B4) are equipped with a well-formed ciliary band and properly formed mouth, developing post-oral arms, and elongated aboral ectoderm leading to a very well-formed vertex (tip of the aboral ectoderm). On the contrary, TRIM treated larvae (Fig. 4 B4, B5, B6) are smaller in size, with a prismatic shape suggesting defects in ectodermal patterning guiding skeletal growth as well as skeletal growth itself. TRIM treated larvae do not have a properly formed ciliary band, while the oral ectoderm surface appears flattened compared to controls.

Nonetheless, a mouth opening implies that the gene regulatory network controlling this process is not affected by Nos inhibition. The presence of an open mouth in Nos deficient larvae suggests that the phenotype in question is not a result of drug toxicity or developmental arrest since it has been demonstrated that mouth occurs only at the larval stage, implying that the prismatic looking trim treated specimens are larvae and not an embryo in which developmental process has been blocked. Interestingly it seems that NO is not necessary for mouth development compared to its role in Amphioxus (Annona et al. 2017).

FISH of genes labeling the affected cell types after Nos inhibition validated the observation based on the larval morphology (Fig. 4.25). Double FISH for the ciliary band marker, identified in this thesis, *Sp-Fbsl_2*, and the lower oral ectoderm marker and the cells around the anus *Sp-Bra* showed that the expression of both genes is altered in trim treated larvae. In more detail, *Sp-Fbsl_2,* which is expressed in all the cells of the ciliary band of control larvae, is found to be expressed in cell patches in the larvae tread with TRIM, while both domains of brachyury show a slight expansion towards more anterior domains. On the contrary, the expression of the upper oral ectoderm marker *Sp-Gsc* appears disorganized in scattered ectodermal cells compared to the observed in control larvae. The overall ectodermal defects are also revealed by FISH of the apical plate marker *Sp-Hbn* and the aboral ectoderm marker *Sp-Spec2a. Sp-Hbn* is typically expressed in a ring of cells surrounding the apical organ, while *Sp-Spec2a* is uniformly found in the larva's aboral ectoderm. *Sp-Hbn* is expressed in a small subset of cells surrounding the apical organ in TRIM treated

larvae in a formation that does not resemble a ring, while the expression of *Sp-Spec2a* appears to have expanded towards the anterior domain. Note the smaller subset of *Sp-Spec2a* negative cells in TRIM treated larvae compared to controls.

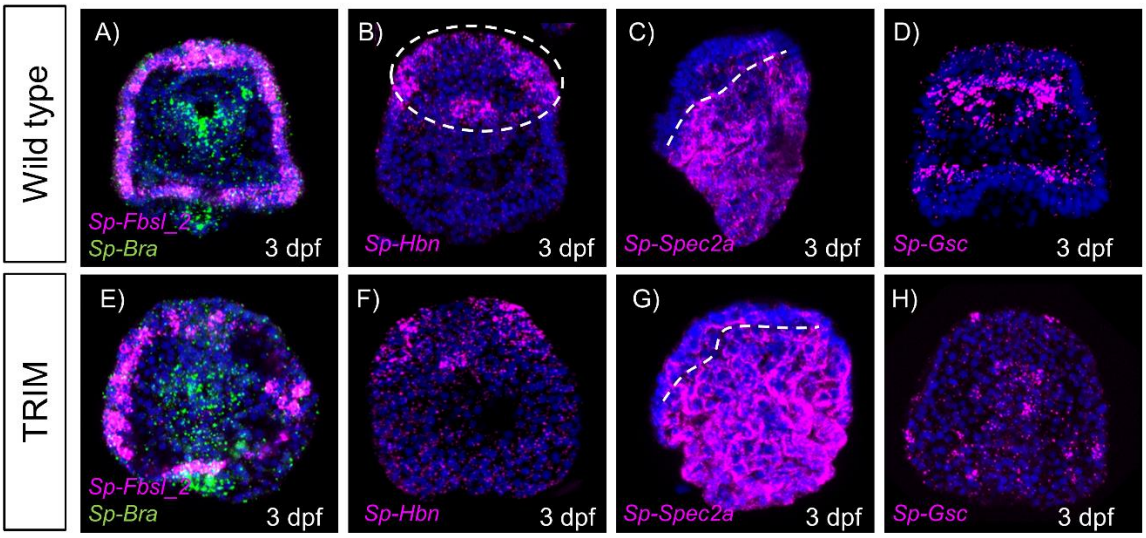


Figure 4.25. Molecular characterization of NO deficient 3 dpf larvae. FISH using specific antisense probes for *Sp-Fbsl_2 and Sp-Bra, Sp-Hbn, Sp-Spec2a,* and *Sp-Gsc* in wild type (A-D) TRIM treated (E-H) larvae. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. The dotted circle highlights the apical plate borders, as indicated by the expression of Sp-Hbn. The dotted line highlights the borders between the Spec2a positive and negative domains.

Overall, the data regarding the morphology and molecular signature of the Nos deficient larvae suggest that NO is essential for proper morphogenesis and larval development, whereas it is not essential for mouth formation. Knowing which are the putative cell types in which NO is detected and possibly exhibits its function, next, I plotted the average expression of genes involved in the canonical pathways through which NO operates (Fig. 4.26).

According to the scRNA-seq analysis, transcripts for MAPK, PI3-K, and JNK kinases were found to be differentially expressed in the apical plate, upper oral ectoderm (domain in which *Sp-Gsc* is expressed), and the neuronal putative broad cell types, while genes involved in the traditional *sGC* and *cGMP* mediated pathway were found to be enriched in the apical plate, esophageal muscles, anal, intestinal, pyloric and cardiac sphincters, and esophageal cells. The presence of the canonical pathway components in the affected by *Nos* inhibition cell types in the apical plate and upper oral ectoderm suggests that NO regulates ectodermal morphogenesis through the canonical pathway. Moreover, the enrichment of the same components

in the cell types in which *NosA* and *B* are expressed and thus NO is produced (neurons, anal, pyloric and cardiac sphincter, and esophageal cells) implies that NO-mediated signaling is activated in the cells that are producing it or that NO produced by neighboring cells is able of initiating a signaling cascade to other NO producing cells. According to the single cell transcriptomic analysis, the canonical signaling cascade can be initiated in the intestinal cells, possibly responding to the NO produced by the NosA positive anal and pyloric sphincter neighboring cells.

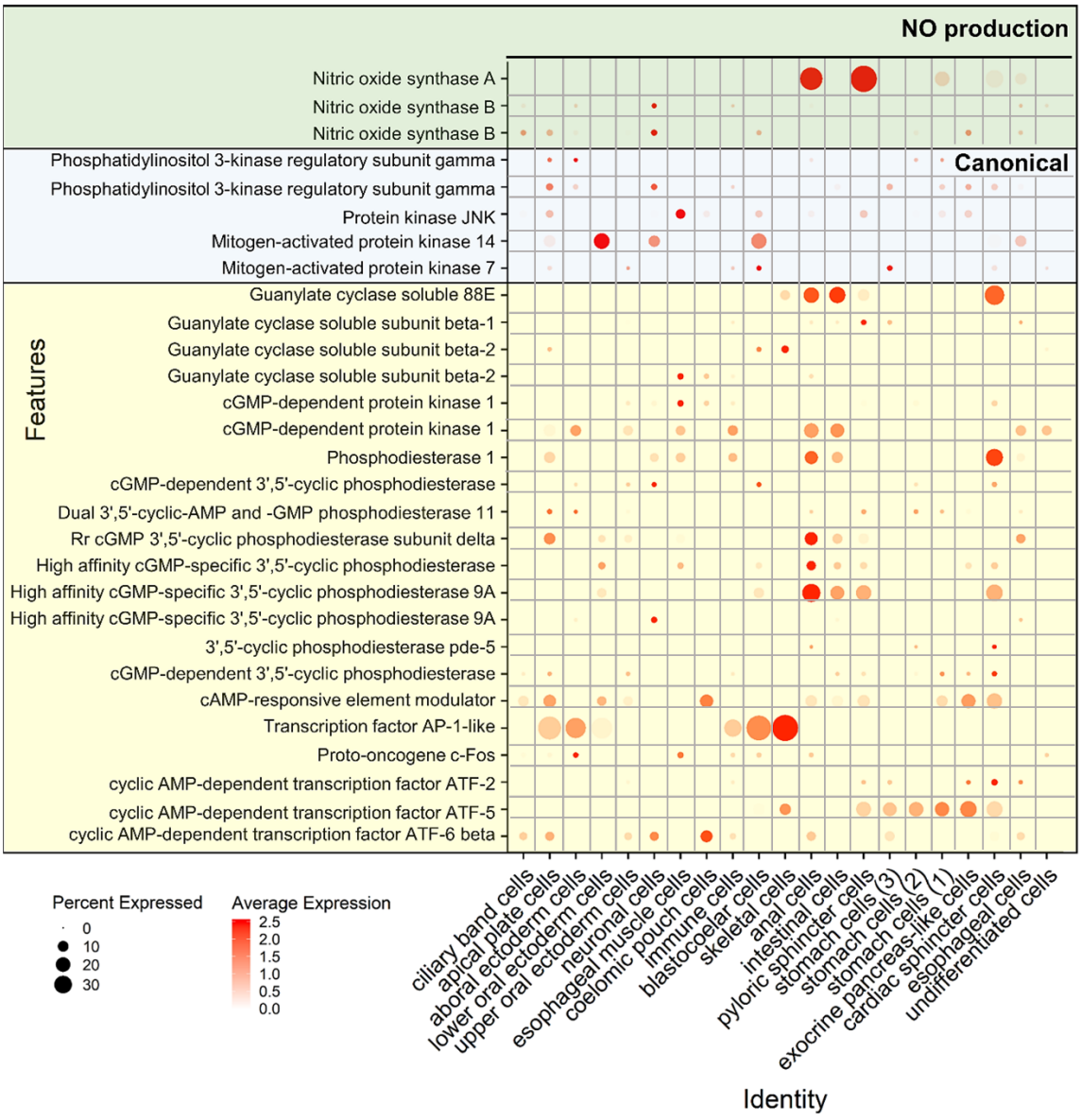


Figure 4.26. NO-mediated canonical signaling pathway in the sea urchin larva. Dotplot showing the average expression of nitric oxide synthases and genes involved in the canonical signaling cascade. Signaling activated by MAPK, PI3-K, and JNK is highlighted in light blue, while the one mediated by sGC is shown in yellow.

One important observation from a transcriptomic perspective is that the machinery involved in the canonical pathway is present in the esophageal muscle, pyloric, and cardiac sphincter cell types suggesting that they can respond to NO. Such a response is in line with the observations in *H. pulcherrimus* in which NO has been implicated with the relaxation of esophageal muscles, cardiac and pyloric sphincter cells (Yaguchi and Yaguchi 2019) of the larva and suggests a similar role also in *S. purpuratus.*

There is no transcriptomic evidence concerning the ciliary band and aboral ectoderm cell types that are severely affected by Nos inhibition suggesting that the normal NO regulation occurs through the canonical signaling in these cells. In the case of ciliary band, the identified presence of NosB and NO, as indicated in this study, suggests that NO is present and possibly regulating the ciliary band formation process possibly through the non-canonical signaling: S-nitrosylation. However, future studies will be needed to test this hypothesis. Concerning the mechanism by which NO might be regulating the aboral ectoderm morphogenesis at this developmental stage, there is no evidence supporting the expression of any Nos gene or NO production. Notably, scRNA-seq, FISH, and IHC suggest that *Sp-NosA* is expressed in the aboral ectoderm at 2 dpf gastrula, which is also the time-point the larvae were exposed to TRIM, allowing one to speculate that this early expression of *NosA* might explain the phenotype observed in the aboral ectoderm formation upon NO production inhibition. Whether the inhibition of Sp-NosA function and thus NO production at gastrula stage is responsible for the aboral ectoderm malformations observed in Nos deficient larvae remains to be verified.

4.2.7. Neurotransmitter release

The last part of the neuronal signaling section focuses on the machinery sea urchin neurons utilize to translocate and release synaptic vesicles containing conventional neuromodulators. Synaptic vesicles loaded with neuromodulators are released into the synaptic cleft via regulated exocytosis, triggered by an action potential arriving in the axon terminal that activates voltage-gated channels.

The vertebrate neurotransmission machinery includes membrane attachment, trafficking, and docking proteins such as syntaxins, required for intracellular vesicle trafficking and secretion. On the other hand, vamps, synaptobrevins, synaptophysin,

SNAPs, snapin, and tomosyn proteins are involved in the synaptic vesicle fusion with the plasma membrane. In contrast, NSF-ATPase ensures the dissemble of the vamps, synaptobrevins, and SNAPs complex once vesicle fusion is complete (Becher et al. 1999, Teng et al. 2001, Tian et al. 2005, Chen et al. 2011). Synaptotagmins are anchored proteins on the synaptic vesicle membrane, acting as a calcium sensor, regulating the release of the synaptic vesicle cargo (Fernandez-Chacon et al. 2001), whereas neurexins ensure neurons stay connected in the synapsis (Geppert et al. 1992, Ushkaryov et al. 1992). Clathrin is essential for coating and shaping of the vesicles, and thus intercellular trafficking, while clathrin-coated vesicles are recognized by the plasma membrane through interaction with the adaptor protein (AP-2) complex, promoting endocytosis of cargo (Boucrot et al. 2010, Kirchhausen et al. 2014). Finally, the amphiphysin/dynamin complex has also been associated with endocytosis of synaptic vesicles (Takei et al. 1999).

Genes involved in the synaptic signaling mechanism have been found in all metazoans, even protists (Burkhardt and Sprecher 2017). The ancestral role of this system remains unknown, although it is possible it has been utilized as a form of cell to cell communication. In the sea urchin, such neurotransmission signaling components have been identified at a genomic level, although their specific role and spatial localization remain unknown (Burke et al. 2006). The most studied gene involved in sea urchin neurotransmission is Synaptotagmin I (SynB), and it has been broadly used as an echinoderm pan-neuronal marker (Burke et al. 2006).

To expand our current knowledge on the neurotransmission machinery operating in the sea urchin, I interrogated the data for the genes mentioned above and plotted their average expression (Fig. 4.27 A). Interestingly, all of the known genes involved in neurotransmission were found differentially transcribed in the neuronal putative broad cell type of the 3 dpf pluteus larva, which is indicative of a similar role as in more complex nervous systems and allowed me to provide a provisional overview of the neurotransmission machinery (Fig. 4.28 B). Surprisingly, genes involved in synaptic signaling were expressed in almost all the cell types with a significant enrichment across the digestive tract ones, which suggests that this trafficking and secretion system has been widely spread and co-opted by many cell types even within the same organism.

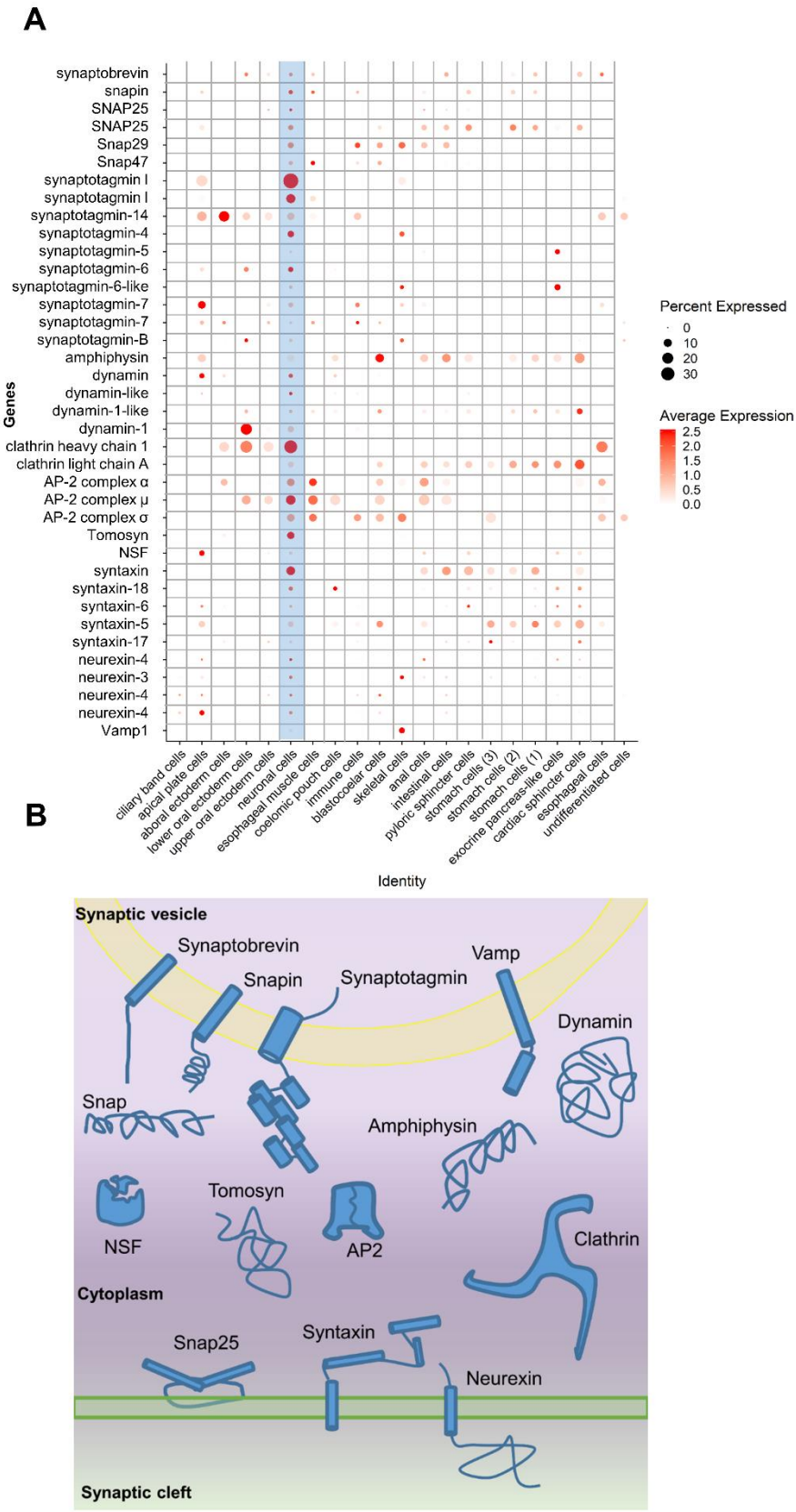


Figure 4.27. Neurotransmission mechanism components. A) Dotplot showing the average expression of genes involved in different aspects of synaptic vesicle trafficking, docking, and release. B) Schematic representation of the putative sea urchin neurotransmitter machinery.

4.3. Patterning and evolution of anterior neuroectoderm

The scRNA-seq analysis discussed in this chapter revealed unprecedented neuronal diversity at gastrula and early pluteus developmental stages, identifying 6 and 12 distinct neuronal populations, respectively. Amongst those neuronal populations, the striking majority of the gastrula (5 out of 6) and almost half of the pluteus ones (5 out of 12) were associated with anterior neuroectoderm (ANE) and apical plate (AP) domains (Figs. 4.8, 4.11). Based on the observation regarding the high number of neuronal types originating from the ANE and AP, my attention was focused on investigating how these two cell types can give rise to such a high neuronal complexity.

As previously mentioned in Chapter 1, ANE is one of the main neurogenic regions of the embryo giving rise to distinct neuronal types. For instance, 2-4 bilateral pairs of serotonergic neurons arise from this domain at gastrula stage. At larval stages, this domain appears to be flattened and thus called the apical plate. A thickened epithelium residing in the center of the apical plate contains sensory neurons and is commonly referred to as apical organ. Apical organs are phylogenetically widely spread to various marine metazoan larvae, including protostome (e.g., mollusks and annelids) and deuterostome larvae (e.g., echinoderms, hemichordates, and cephalochordates), and they are found in the most anterior ectodermal part of the larva. Apical organs are present during the larval stages and disappear after metamorphosis, suggesting that apical organs are used to detect environmental cues promoting settlement and metamorphosis. Apical organs are equipped with longer sensory cilia than the rest of the larval cilia, known as apical tuft and diverse neuronal populations.

For animals in which molecular data is available, spatial gene expression data have been used to compare the molecular signature and especially the transcription factor content of apical organs and the neuronal distribution within them. This lead to the formation of comparative gene expression maps showing conserved co-expression gene expression patterns. In a study by Marlow and coworkers, it was demonstrated that the transcription factors *Six3* and *Foxq2* are essential for the apical organ formation and patterning in several animal groups (cnidarians, annelids, hemichordates, echinoderms, and cephalochordates (Marlow et al. 2014)). Moreover, it was shown that in the same animal groups, the apical organ's

restriction to the most anterior domain is a result of Wnt signaling inhibition in that domain. Furthermore, the neuronal composition of the apical organs in cnidarians, mollusks, annelids, hemichordates, echinoderms, and cephalochordates is very similar harboring in all cases neurons that produce the neurotransmitter serotonin (Marlow et al. 2014). Whether the serotonergic neurons' formation by apical organ cells is related to the conserved gene regulatory wiring is unknown. Nonetheless, the current data suggest that the apical domain wiring predates bilaterians.

Additional comparative studies demonstrated that elements of the gene regulatory signature of the apical organs are also present in bilaterians with a direct mode of development and are found to be utilized in wiring anterior brain structures, suggesting a possible conservation of the anterior gene regulatory wiring and its co-option into controlling the formation of anterior brain tissues (Loosli et al. 1998, Steinmetz et al. 2010, Posnien et al. 2011). Mainly, it has been hypothesized that apical organs and the forebrain share a common evolutionary history. The conserved role of the transcription factors *Rx, Six3, Foxq2*, and *Fez* in the specification of both apical organs and forebrain combined with the multifunctional role of both apical organs and forebrain regions in hormonal secretion, non-visual light perception, and circadian rhythm regulation are some of the main points supporting this hypothesis (Tosches and Arendt 2013, Marlow et al. 2014).

Previous studies in sea urchin embryos and larvae demonstrated that the apical plate region shows compartmentalization of gene expression to distinct apical plate domains such as distal, medial, and central. Since the scRNA-seq analysis predicts only one anterior neuroectoderm apical plate cell type, subclustering analysis was applied in a similar approach as performed to assess neuronal diversity. To identify the different anterior/apical domains, the average expression of genes involved in restricting the size of the ANE/AP, genes associated with the apical tuft, and transcription factors known to be expressed in distinct such domains were plotted.

Subclustering of the 2 dpf gastrula anterior neuroectoderm and of the 3 dpf apical plate cell clusters with using the resolution of 1.0 resulted in the generation of 4 clusters at both developmental stages, which I was able to map based on the co-expression or lack of it of known gene markers (Fig. 4.28).

Antagonists of Wnt signaling such as *Frizzled* and *Dkk* genes shown to be essential for the restriction of the anterior neuroectoderm and thus apical plate to the most

anterior part of the embryo/larva are found to be enriched in two gastrula ANE subpopulations (ANE1 and ANE3), while the gene *ankyrin-containing gene-specific for apical tuft-1(Ankat-1*) (Yaguchi et al. 2010), which is essential for the apical tuft formation is found in all ANE subpopulations except from ANE3 (Fig. 4.28 A). On the other hand, the enzyme glutathione S-transferase theta-1 (*Gsst*) involved in the regulation of mechanical reception and swimming behavior (Jin et al. 2013) is found in all of the ANE subtypes (Fig. 4.28 A). Transcription factors such as *Sp-Nkx2.1, Sp-Fez, Sp-Rx* (Takacs et al. 2004, Yaguchi et al. 2011, Valencia et al. 2019) are highly correlated with the central anterior neuroectoderm domains as demonstrated by previous studies and are all found to be enriched in the ANE 3 population (Fig. 4.28 A). Moreover, the combinatorial expression of genes that are shown to be expressed in the distal ANE and not within the central ANE domain such as *Sp-Hbn*, *Sp-Six3, Sp-Tbx2/3*, and *Sp-FoxG* (Wei et al. 2009, Yaguchi et al. 2011, Valencia et al. 2019) is detected in the ANE 1 population, while the transcription factor *Sp-Hlf* marker of the central and dorsal anterior neuroectoderm domain (Howard-Ashby et al. 2006) is found enriched in populations ANE 2 and 3 (Fig. 4.28 A). Concerning the subpopulations of apical plate present in the 3 dpf pluteus stage, plotting the same markers resulted in similar results. In more detail, the Wnt signaling antagonism was found to be active also in two apical plate populations (AP3 and AP4), the apical tuft associated gene *Sp-Ankat-1* was found present in two subtypes (AP2 and AP3), while *Sp-Gsst* transcripts were ubiquitously expressed in all of the AP subpopulations (Fig. 4.28 B).

Similarly to the 2 dpf dataset analysis, *Sp-Nkx2.1, Sp-Fez,* and *Sp-Rx* were found enriched in the apical plate 1 subtype, while *Sp-Hbn, Sp-Six3,* and *Sp-Tbx2/3* were differentially expressed in the apical plate 4 subtypes (Fig. 4.28 B). Notably, transcripts of most of the already described gene markers were found to be enriched in distinct ANE/AP domains, in line with the available from literature information, except from the transcription factor *Sp-Foxq2*. As mentioned above, *foxq2* is one of the key regulators of the apical organ formation and its role is evolutionary conserved in several animal groups (cnidarians, annelids, hemichordates, echinoderms and cephalochordates). The reason for the absence of *Foxq2* transcripts from all of the produced datasets is purely technical and is based on the reference transcriptome used for this analysis that lacks *Foxq2* transcripts.

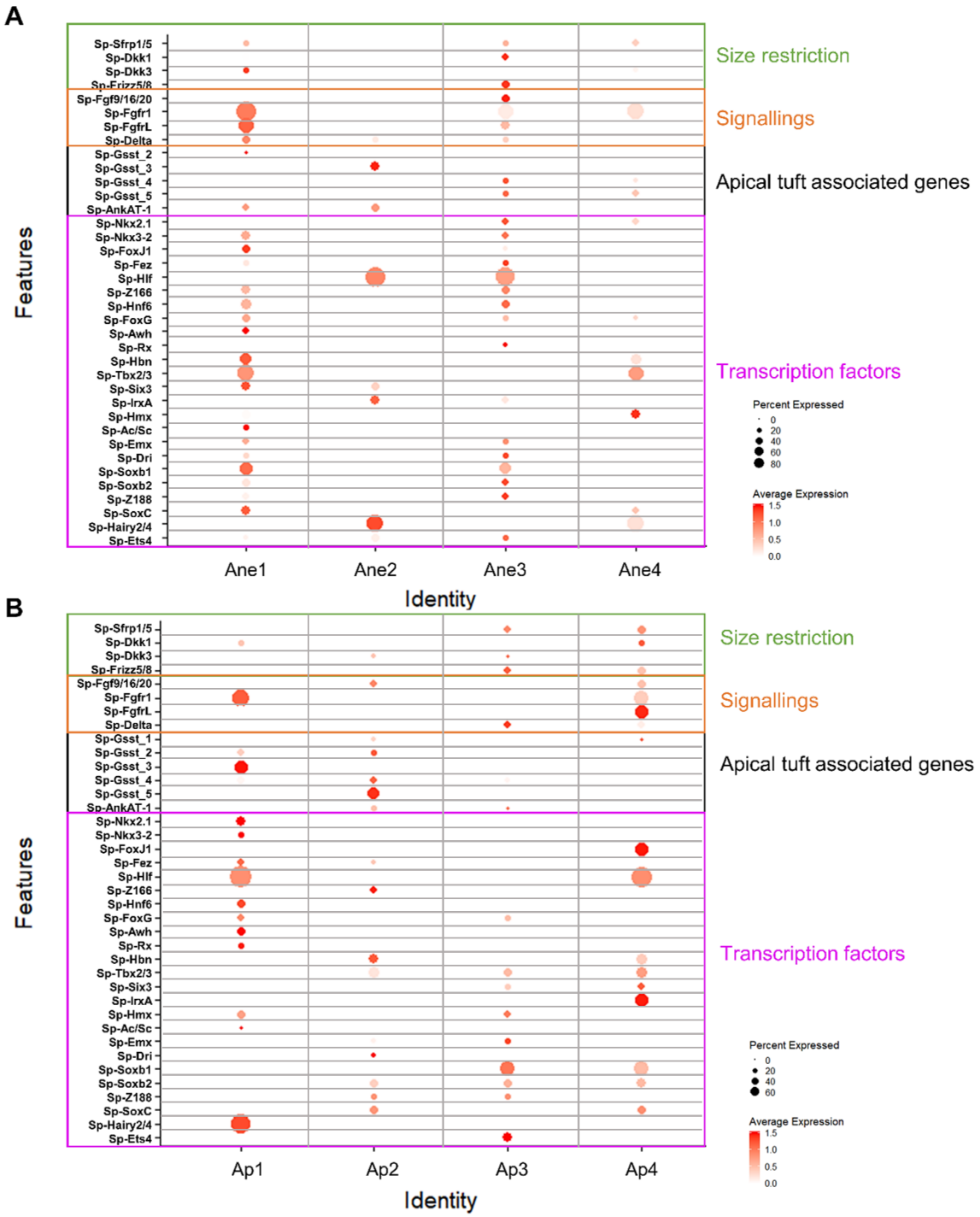


Figure 4.28. Anterior neuroectoderm and apical plate gene markers. Dotplots showing the average expression of genes involved in ANE/AP size restriction, signaling pathways, apical tuft formation, and function as well as key transcription factors at gastrula (A) and early pluteus (B) developmental stages.

Next, I performed FISH and IHC of several of the analyzed gene markers to validate *in vivo* the co-expression predictions deriving from the scRNA-seq analysis (Fig. 4.29 B). In order to also understand the subpopulations in which *Foxq2* is

expressed, I cloned the *Sp-Foxq2* gene and performed FISH and IHC with gene markers that were included in the single cell analysis.

FISH performed on gastrulae using antisense RNA probes for *Sp-Hbn* and *Sp-Emx* showed that *Sp-Hbn* (Fig. 4.29 A1) and *Sp-Emx* (Fig. 4.29 A3) transcripts are co-localized in the distal ANE region, while *Sp-Emx* transcripts are also found within central ANE. *Sp-Frizz5/8* (Fig. 4.29 A2) is expressed in the central and medial parts of the ANE as well as *Sp-SoxC* (Fig. 4.29 A4). Co-expression analysis of the Fgf ligand-receptor *Sp-Fgfr1* and the transcription factors *Sp-Nkx2.1* and *Sp-Foxq2* showed that their transcripts co-localized in the central ANE domain (Fig. 4.29 A5, A6). Since, according to the scRNA-seq analysis, *Sp-Nkx2.1* and *Sp-Fgfr1* partially co-localize only in the ANE population 3 (Fig. 4.28 A), and given that *Sp-Foxq2* transcripts are found to col-localize with transcripts of both genes, it is safe to say that also Sp-Foxq2 is expressed in the ANE 3 domain. In support of this, FISH paired with IHC shows that there are no *Sp-Foxq2* positive cells that do not express Sp-Nkx2.1 (Fig. 4.29 A5, A6) and that all *Sp-Foxq2* positive cells produce *Sp-Fgfr1* (Fig. 4.29 A5). Furthermore, in the same ANE population, partial co-localization of *Sp-IrxA* and *Sp-Fgf9/16/20* was found (Fig. 4.29 A7, A8, A9, A11). *Sp-SoxB2* was found to be expressed in the distal and central domains of the ANE (Fig. 4.29 A9), while *Sp-Tbx2/3* transcripts were detected in the distal and medial ANE region (Fig. 4.29 A12).

Performing FISH experiments on 3 dpf pluteus larvae for *Sp-Six3, Sp-Tbx2/3* and *Sp-Hbn* and *Sp-IrxA* showed a broad expression of these genes in the apical plate region, although all three of them were found to be expressed in the central region of it (Fig. 4.29 B1, B2, B3, B8). The partial co-localization of *Sp-Fgf9/16/20* observed at 2 dpf gastrula seems to be abolished at pluteus stage, since no double-positive cells for such genes could be detected (Fig. 4.29 B5). *Sp-Frizz5/8* was found to be expressed in the central AP domain, while *Sp-Emx* transcripts were detected in the most distal AP regions (Fig. 4.29 B4, B6). Concerning the expression of *Sp-Foxq2*, it was found to be co-expressed in the same cells are *Sp-Nkx2.1* (Fig. 4.29 B7), similar to what has been observed at gastrula stage, suggesting that *Sp-Foxq2* is expressed by the AP1 cell population.

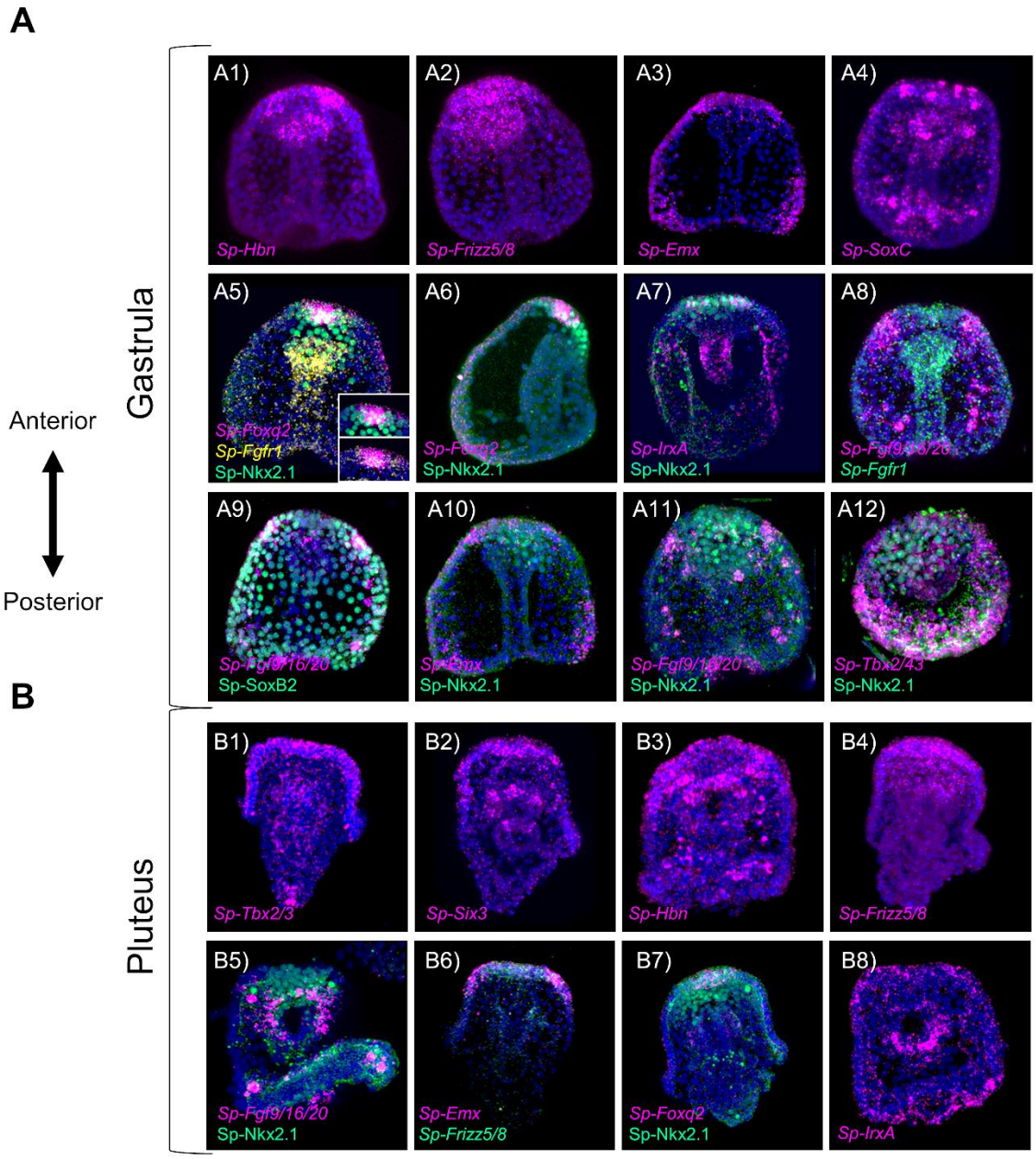


Figure 4.29. Molecular signature of ANE/AP regions. A) FISH of *S. purpuratus* 2 dpf embryos with specific antisense probes transcription factors *Sp-Hbn* (A1), *Sp-Frizz5/8* (A2), *Sp-Emx* (A3, A10), *Sp-SoxC* (A4), *Sp-Foxq2* (A5, A6), *Sp-Fgfr1* (A5), *Sp-IrxA* (A7), *Sp-Fgf9/16/20* (A6, A9, A11) and *Sp-Tbx2/3 (*A12*).* B) FISH of *S. purpuratus* 3 dpf larvae with specific antisense probes transcription factors *Sp-Tbx2/3* (B1), *Sp-Six3* (B2), *Sp-Hbn* (B3), *Sp-Frizz5/8* (B4, B6), *Sp-fgf9/16/20* (B5), *Sp-Emx* (B6), *Sp-Foxq2* (B7), *Sp-IrxA* (B8)*.* FISH shown in figures A5-A7, A10-A12, B5, and B7 are paired with immunohistochemical detection of the transcription factor Sp-Nkx2.1, while FISH shown in A9 is paired with IHC detection of the transcription factor Sp-SoxB2. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections.

Combination of the information obtained by the scRNA-seq analysis and the spatial/temporal expression of the aforementioned gene markers, I was able to build a provisional map showing the spatial distribution of the different anterior

neuroectoderm and apical plate regions (Fig. 4.30 A, B). Taking into account the possible high transcriptomic similarity between the subpopulations of the anterior neuroectoderm and apical plate putative broad cell types, I also repeated the subclustering analysis keeping all the variables constant and by changing the FindClusters resolution from 1.0 to 1.5. Since both anterior neuroectoderm and apical plate clusters consist of relatively similar number of cells (Table 3.1), in case both datasets consist of the same subpopulations changing the resolution would evenly increase the number of sub-clusters produced. This analysis led to the generation of 6 anterior neuroectoderm subtypes and 10 apical plate cell types, suggesting using resolution 1.0 could not identify subtle transcriptomic differences between the subpopulations (Fig. 4.30 C). These preliminary data suggest that there is an increasing complexity in apical plate patterning and wiring, which could be in line with the increasing demand for neurogenesis that takes place during larval development. Overall the results of the initial subclustering analysis depict a core of four distinct regulatory domains at both 2 and 3 dpf developmental stages. Whether the additional cell types occurring while increasing the resolution reflect additional cell types or different regulatory states of the already identified anterior/apical domain subpopulation will be determined by future studies.

In a review study by Beccari and coworkers, the authors used molecular data deriving from studies in Xenopus, Medaka, Zebrafish, chick, and mouse to reconstruct the backbone of a conserved gene regulatory network operating during the specification of the anterior neural plate in vertebrates (Beccari et al. 2013). This GRN consists of gene homologs with conserved interactions among those species that are all necessary for the specification of the anterior neural plate that will give rise to the forebrain and the distinct structures telencephalon, hypothalamus, retinal field, and diencephalon (Figure 4.32 A). Taking into account the information provided by the former study, the hypothesis that the molecular signature of the apical organs contains evolutionary conserved elements that predate bilaterians, and that the apical organs share molecular and functional features with forebrain regions, I set out to investigate the extent to which these elements are conserved in the sea urchin.

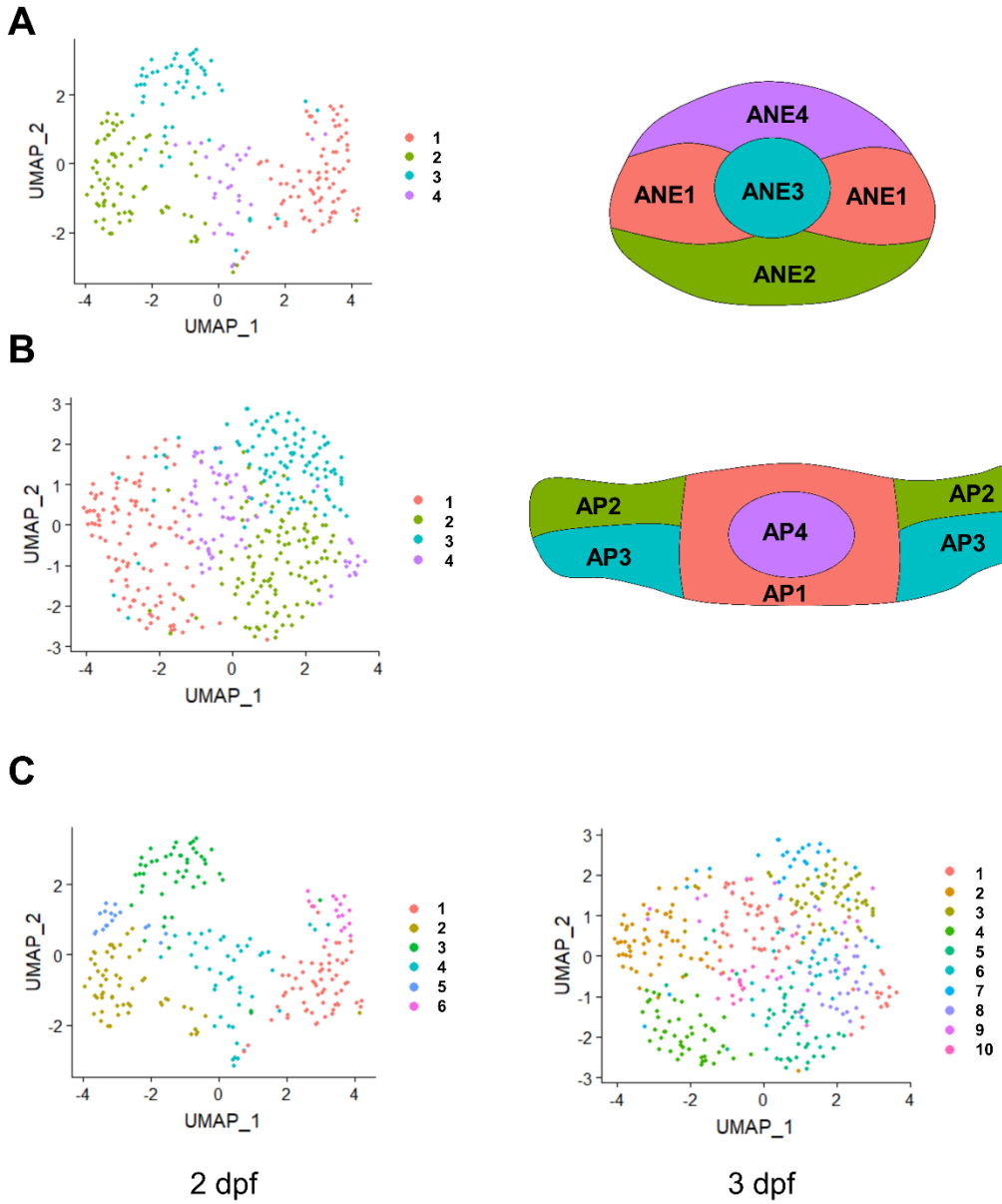


Figure 4.30. Mapping of the distinct ANE/AP regions. A) UMAP showing the annotated anterior neuroectoderm populations of 2 dpf gastrula embryo and their putative spatial location on the ANE domain. B) UMAP showing the annotated apical plate populations of 3 dpf pluteus larva and their putative spatial location on the AP domain. C) UMAP showing additional subclustering analysis of the anterior neuroectoderm and apical plate putative broad cell types. Note: The coloring of the ANE and AP schematic representations in A and B corresponds to their clustering color-code and does not imply a similarity between the ANE and AP domains.

Plotting of the sea urchin homologs of the genes involved in the anterior neural plate specification led to identifying a conserved molecular signature between the sea urchin apical plate and the vertebrate anterior neural plate (Fig. 4.31). In more detail, 15 homologous genes including the transcription factors *Sox2, Otx2, Fez, Six3/6, Prox1, Rx, Nkx2.1, Lhx2, Pax6, Tcf3/4/12-like, Mab21l2* and *Lmo4* as well as the Fgf and Sonic Hedgehog ligand Fgf8/17/18 and Shh respectively, were found in

apical domain cell types. Notably, all of the genes mentioned above were found to be expressed in anterior neuroectoderm regions of the 2 dpf gastrula (Fig. 4.31 A), apical plate domains of the 3 dpf early pluteus larva (Fig. 4.31 B) as well as in the apical regions originating from the subclustering of the 2 and 3 dpf integrated data (Fig. 4.31 C). FISH experiments for the genes *Prox1, Pax6, Six3/6, Otx2,* revealed the spatial localization of those transcripts in the apical plate domain of the 3 dpf early pluteus in line with the single cell predictions (Fig. 4.31 D).

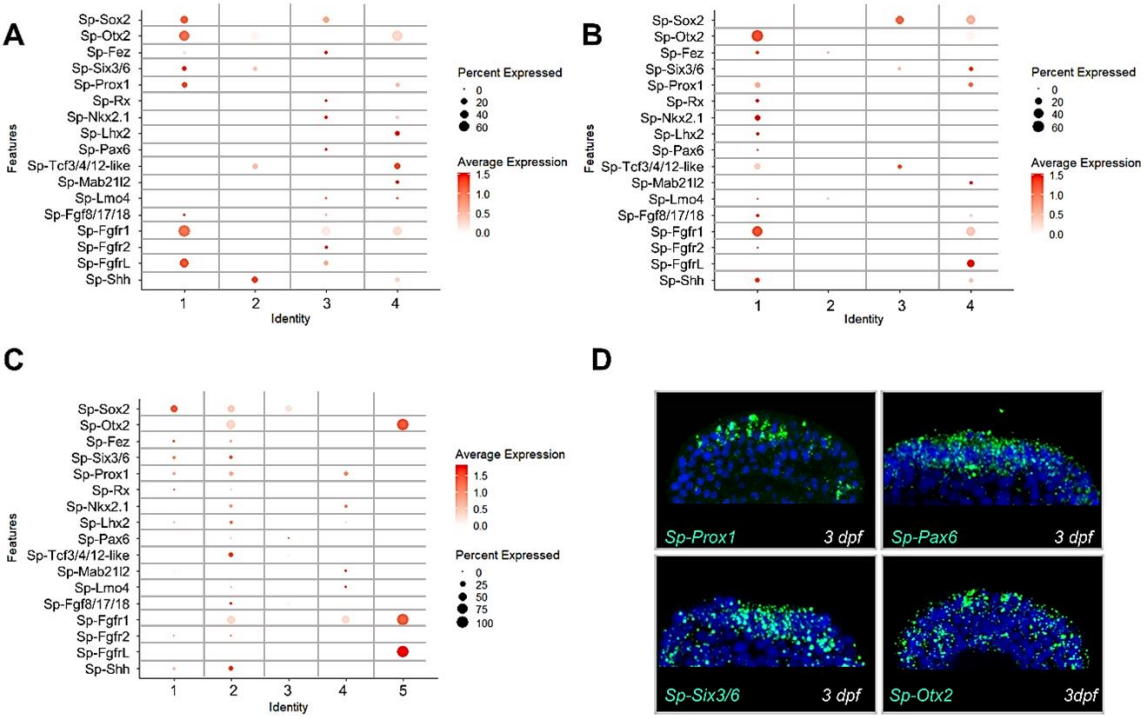


Figure 4.31. The anterior neural plate gene toolkit is conserved in sea urchin. Dotplots showing the average expression of transcription factors and signaling molecules involved in the specification of the anterior neural plate in vertebrates at 2 dpf (A), 3 dpf (B), and 2 & 3 dpf integrated (C) datasets. D) FISH using antisense RNA probes for the *Sp-Prox1*, *Sp-Pax6*, *Sp-Six3/6* and *Sp-Otx2* genes. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. Note: the vertebrate homolog name was used.

These results provide insight into the conservation of the molecular toolkit used to wire these two anterior domains as they provide evidence that the apical domain and anterior neural plate specification pathways are occurring through similar molecular toolkits (Figure 4.32 B). Moreover, these data are in line with the hypothesized evolutionary origins of the forebrain from the invertebrate apical plates. However, no definite conclusions can be drawn until similar toolkits are identified in a broader spectrum of organisms. Additionally, future studies are needed to identify the wiring between those conserved genes to understand whether

the conservation observed at a patterning level is also reflected at a gene regulatory network level.

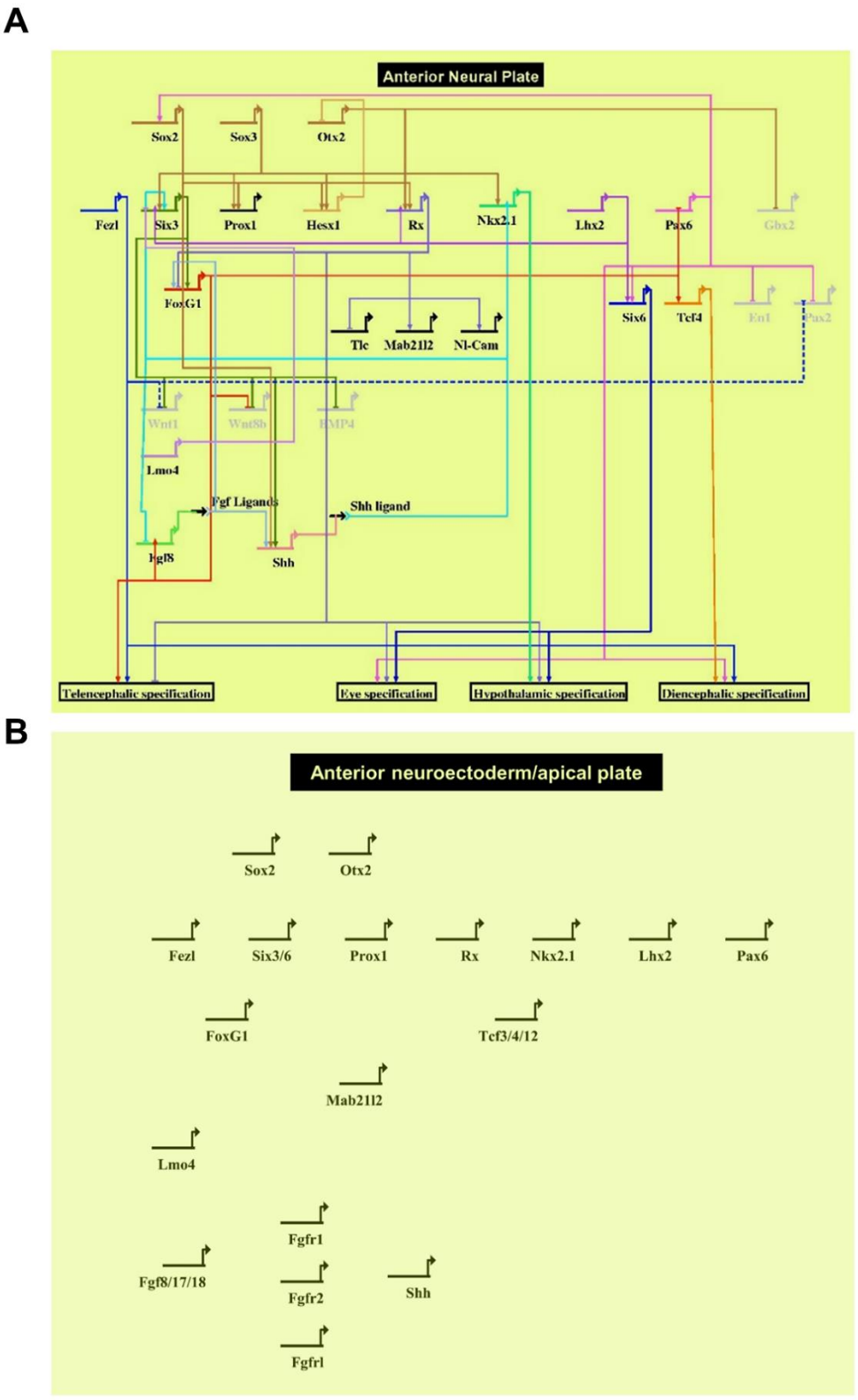


Figure 4.32. Apical neural plate and ANE/AP molecular signature conservation. A) Backbone of a GRN guiding the specification of apical neural plate adapted from Beccari, Marco-Ferreres, et al. 2013. B) Sea urchin homologs of genes shown in A with yet unknown wiring as identified by this work. Note: the vertebrate homolog name was used.

Chapter 5

Pancreatic-like cell types in sea urchin

Previous studies identified two pancreatic-like cell populations residing within the sea urchin larva's digestive tract that bear exocrine and endocrine-like characteristics and are involved in the feeding process. Based on these observations, I continue my analysis investigating the pancreatic signature of such cell types and whether other cell types also utilize a similar pancreatic-like molecular program. This chapter contains the results of this analysis identifying neuronal and non-neuronal cells with an endocrine pancreas-like signature and at least one neurosecretory type, whose genetic wiring is very similar to the vertebrate β-endocrine cells. Detailed results discussed in this chapter can be found in the Non-book component of the thesis.

5.1. Pancreatic cell types

The survival of an organism depends on its capability to find and catch food and digest the food source. Digestion relies on the production and secretion of digestive enzymes that metabolize large biomolecules like proteins to amino acids and hormones involved in the metabolism and glucose homeostasis. In vertebrates, these processes are being orchestrated by a multifunctional organ, the pancreas.

The pancreas is a complex organ that bears specialized cell types that are responsible for its multi-functionality. In detail, pancreas is a mixed endocrine and exocrine gland consisting of exocrine and endocrine cells producing digestion zymogens (carboxypeptidases, amylases, lipases, proteinases) and hormones (insulin, glucagon) that are in direct contact with a ductal epithelium that neutralizes those enzymes. The cellular composition of pancreas is widely conserved among vertebrates, although the spatial distribution varies across taxa.

For instance, in the mammalian pancreas, the endocrine cells cluster together to form islets (islets of Langerhans), which are embedded within the exocrine tissue, and that cell arrangement is observed only in the mammalian clade (Fig. 5.1).

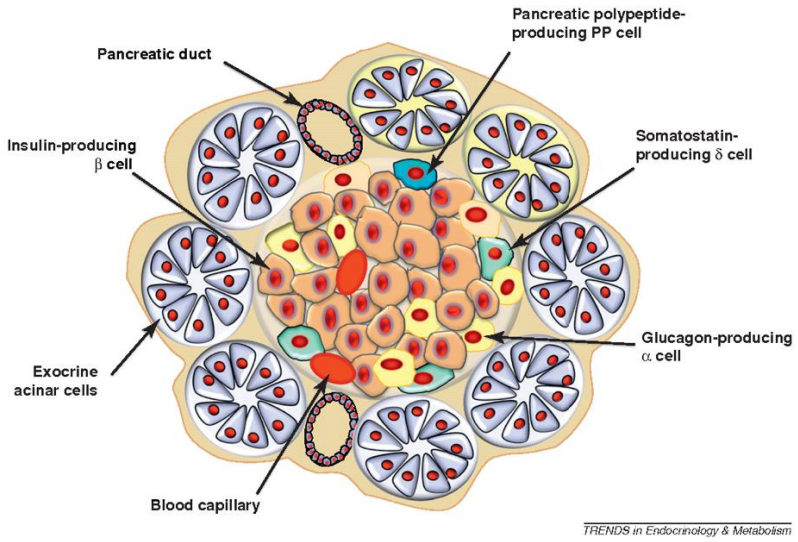


Figure 5.1. Pancreatic cell types in mammals. Schematic representation of a mammalian pancreas depicting the embedding of the five distinct endocrine cell types within the exocrine acinar tissue (Figure adapted from Efrat and Russ 2012).

In addition, the cells constituting the exocrine tissue are grouped into distinct structures, the acini (acinar cells). The mammalian islets of Langerhans consist of five specialized cell types, including β-cells that produce the hormone insulin that

maintains the bloodstream glucose at low levels, α-cells that produce glucagon with an opposite effect in regards to insulin, δ-cells that produce and secrete somatostatin regulating the release of both insulin and glucagon. In contrast, PP and ε-cells producing the pancreatic polypeptide and hormone ghrelin respectively, are involved in food intake regulation (Edlund 2002, Batterham et al. 2003, Rindi et al. 2004, Rindi et al. 2004).

In other vertebrates such as teleosts, the organization of the pancreatic cell types appears to be more simplified with pancreas-like organs consisting of endocrine cells organized in islets but not embedded within the exocrine tissue (Falkmer et al. 1985, Yui and Fujita 1986, Youson et al. 2006). The formation of a distinct pancreas-like organ seems to be a vertebrate innovation as no such structure can be identified in non-vertebrate chordates. However, gene expression studies showed that while no pancreas is present in cephalochordates and tunicates, several pancreatic-like cell types are dispersed throughout the digestive tract (Reinecke and Collet 1998, Olinski et al. 2006, Arda et al. 2013, Lecroisey et al. 2015). In the sea urchin, a non-chordate deuterostome, exocrine pancreas-like cells have been found present in the upper part of the larval stomach organized in a belt-like structure, while endocrine pancreas-like cells producing insulin have been identified in a neighboring region (Perillo and Arnone 2014, Perillo et al. 2016).

Similar, pancreatic-like cell types have been identified in cnidarian members suggesting that a pancreatic-like machinery might have been present in the last common ancestor of cnidarians and bilaterians. In more detail, scRNA-seq of *Nematostella vectensis* revealed genes encoding for secreted digestive enzymes that were restricted to the *FoxA* positive domains of the pharynx cnidoglandular tract (Sebe-Pedros et al. 2018). Furthermore, a similar spatial distribution of exocrine and insulin-producing cells has been reported in ectodermally derived cell types of the sea anemone (Steinmetz et al. 2017).

5.2. Pancreatic development

Organogenesis occurs during embryonic development, and organ-identity is established by GRNs that are promoting distinct regulatory programs on multipotent cells. One of the most well-characterized organogenetic regulatory programs in vertebrates and especially in mice concerns pancreas development (Slack 1995,

Pearl et al. 2009, Wilfinger et al. 2013). The developmental cascade, whose lineage is schematized in Fig. 5.2, includes specific terminal differentiation gene batteries, which members have been identified, and their role has been assessed. Many of these genes are necessary for the specification of the initial endodermal region that will form the pancreas as well as the differentiation of pancreatic progenitors to various pancreatic cell types.

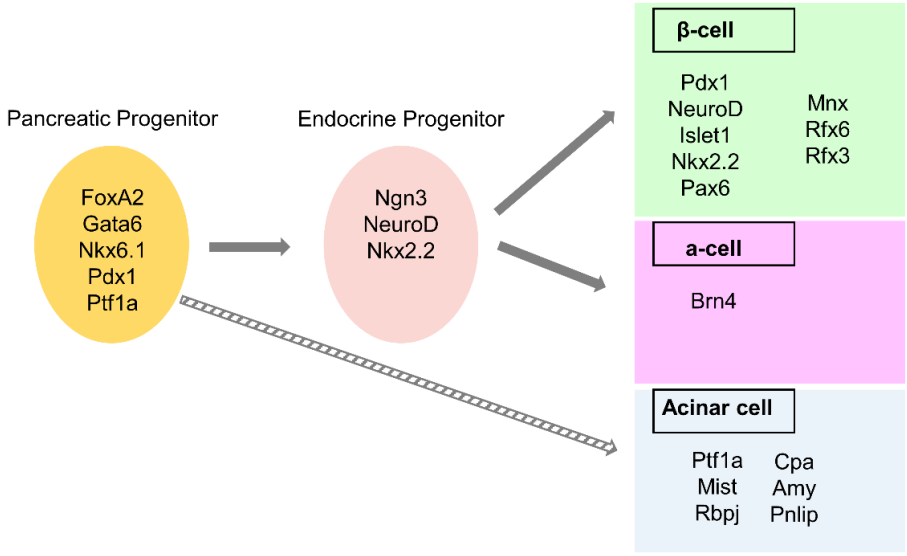


Figure 5.2. Simplified lineage model of pancreatic differentiation cascade in mammals. Key transcription factors that mark each lineage are reported.

In mice, the pancreas develops from protrusions of the foregut after receiving morphogenetic signals (Hedgehog, Fgf, Notch, Wnt, and Tgf-β) from neighboring mesodermal and ectodermal cell types (Edlund 2002, Oliver-Krasinski and Stoffers 2008). The first step of pancreas development is the formation of a multilayered epithelium containing microlumen structures, the pancreatic buds, and is a result of multipotent cell proliferation initiated by signals deriving from the notochord and the surrounding tissues. During the next phase, these microluminar structures fuse to create an epithelial network divided into distinct tip and trunk regions, while in parallel pancreatic cell type specification and differentiation occurs. The specified pancreatic progenitors can give rise to both endocrine progenitors and exocrine acinar cells, while their pluripotent cell state depends on several transcription factors (Grapin-Botton and Melton 2000, Jensen 2004). In particular, *pancreatic and duodenal homeobox 1 (Pdx1)* and *pancreas-specific transcription factor 1a (Ptf1a)* are amongst the first transcription factors to be activated within the primitive gut tube (Burlison et al. 2008), and silencing of either of those genes results in pancreatic

agenesis in mice (Ahlgren et al. 1996, Kawaguchi et al. 2002, Marty-Santos and Cleaver 2016). Both of those transcription factors have been found to have a dual role in the initial steps of pancreatic progenitor specification and cell type differentiation. In particular, *Ptf1a* is also essential for the differentiation of acinar cells (Zecchin et al. 2004), while *Pdx1* is re-utilized to promote β-cells differentiation (Macfarlane et al. 1999, Afelik et al. 2006). The high importance of *Pdx1* in orchestrating the GRN guiding pancreas formation and the endocrine cell fate is highlighted by its early expression necessary to activate, in combination with other transcription factors, a specific regulatory state in the confined region of the primitive gut that will give rise to the pancreas and its later role in the differentiation of the endocrine β-cells.

Subsequent activation of the transcription factors *Ngn3, NeuroD*, and *Nkx2.2* in a subpopulation of pancreatic progenitors seals their fate in differentiating into endocrine progenitors (Schwitzgebel et al. 2000, Dalgin and Prince 2015, Churchill et al. 2017). The process of acquiring an endocrine versus an exocrine fate is being regulated by Notch signaling as it has been demonstrated that active Notch signaling promotes exocrine fate through the activation of *Ptf1a* and the repression of *Ngn3* (Qu et al. 2013). Moreover, the transcription factors *Mist1* and *Rbpj* that are also activated due to Notch signaling regulate exocrine function and promote exocrine fate, respectively (Pin et al. 2001, Fujikura et al. 2007).

The endocrine progenitor cells further differentiate into five types of mature endocrine cells under the influence of specific gene regulatory interactions. In the case of β-cells, their differentiation depends on a gene regulatory core consisting of the transcription factors *Pdx1, NeuroD, Islet1, Nkx2.2, Pax6, Mnx, Rfx6,* and *Rfx3* and silencing of either of those genes results in impairment of β-cell maturation and β-cell function, (Ait-Lounis et al. 2010, Gu et al. 2010, Gosmain et al. 2012, Ediger et al. 2014, Piccand et al. 2014, Pan et al. 2015, Gutierrez et al. 2017). In contrast, the transcription factor *Brn4* is the dominant regulator of glucagon expression in another endocrine cell population, the α-cells (Hussain et al. 2002).

Until recently, identifying the pancreatic cell types and the GRNs underlying their differentiation was solely based on experimental studies spanning from morphological analysis, lineage tracing, and investigation of selected gene candidates to mutagenesis (Arda et al. 2013). Recent advances in single cell transcriptomics have allowed an extensive characterization of the pancreatic differentiation cascade overcoming the limitations of approaches used in the past.

Due to the advantages that this method has to offer, it comes as no surprise that an organ with complex developmental and evolutionary history, such as the pancreas, was amongst the first specimens to be analyzed at a single cell resolution (Baron et al. 2016, Muraro et al. 2016). Nowadays, several scRNA-seq studies have been carried out mostly on human and mice in an attempt to gain insight into the regulatory mechanisms that guide pancreatic development and how disruption of such mechanisms leads to pancreatic diseases (Wang et al. 2016, Wollny et al. 2016, Enge et al. 2017, Tritschler et al. 2017, Ramond et al. 2018, Belle and DeNardo 2019, Gao et al. 2019, Glinka et al. 2019, Luo et al. 2020).

ScRNA-seq technology, apart from allowing the identification of the cell types present in an organism or organ, can also be used as a tool to reconstruct cell type evolution (Arendt et al. 2016). Such an approach was followed in a study by Baron and colleagues to compare the cell types patterning the pancreas of humans and mice at a single cell transcriptomic level (Baron et al. 2016). The findings of this study include the observation of an unprecedented diversity among the inter-species pancreatic cell types, identifying novel types of ductal cells and endocrine progenitors, as well as a high level of molecular signature conservation between human and murine pancreatic cell types (Baron et al. 2016).

5.3. Pancreatic cells and nervous system

Cell types within an organism with similar molecular fingerprints, morphology, and physiological roles usually have shared developmental histories. However, there are cases in which cell types, such as the pancreatic β- cells and neurons, that share many features but have separate developmental origins.

The physiology of both pancreatic β-cells and neurons includes several shared elements (Alpert et al. 1988, Eberhard 2013). For instance, both cell types can produce, secrete and respond to polypeptide hormones (Pearse and Polak 1971, Fujita et al. 1980) and neuromodulators (Reetz et al. 1991, Maechler and Wollheim 1999). Strikingly the mechanism by which β-cells produce, store and secrete insulin upon induction is almost identical to how neurons produce, store and initiate their neuronal signaling (Henquin and Meissner 1984). Moreover, these shared features are also reflected in the repertoire of genes they express and their genetic wiring. The specification of both cell types depends on Hedgehog, Fgf, Notch, Wnt, and Tgf-β signalings (Lasky and Wu 2005, Machon et al. 2007, Belgacem et al. 2016,

Diez Del Corral and Morales 2017, Meyers and Kessler 2017), while transcription factors involved in β-cell differentiation such as *Ngn3, Pdx1, Islet, NeuroD* are also crucial for neural differentiation (Thor et al. 1991, Vult von Steyern et al. 1999, Kim et al. 2001, Schuurmans and Guillemot 2002, Song et al. 2010, Zhu et al. 2017) in mammals.

The many similarities in terms of physiology, function, and molecular signature raises many questions on how these two cell types ended up being similar. Are β-cells and neurons part of the same differentiation lineage? Is the similar gene regulatory program a result of convergent evolution and a co-option of a similar regulatory program? The hypothesis of β-cells and neurons sharing common developmental origin has been discouraged by experimental data showing that these two cell types arise from different germ layers during development. Moreover, inhibition of neurogenesis by removal of the neural tube prior to neural crest migration did not affect the formation of pancreas in rats (Pictet et al. 1976), while transplantation experiments of neural crest cells revealed no contribution in the formation of endocrine cells in chick embryos (Fontaine et al. 1977).

A possible explanation for the genetic similarities of these two cell types is that they arose through convergent evolution by co-option of a neuronal preexisting gene regulatory program (Arntfield and van der Kooy 2011) (Fig. 5.3) . This theory is being supported based on the widespread distribution of neuronal or neuronal-like cells across metazoa that, in many cases, also produce insulin and the reports on lack of specific gut cells producing insulin in several animal clades.

According to the proposed mechanism a mutation on the promoter region of a master gene regulator initiating the neuronal program in neurons allowed its activation in several gut cells and thus promoting a neuronal program (Arntfield and van der Kooy 2011). Whether this initial endocrine cell expressed a common program to all neurons or a program that was present in distinct neuronal populations is not clear. The co-option of a generalized neuronal program is supported by the fact that the nervous system components can produce several pancreatic hormones and that pancreatic cells bear neuron-like features, and that β-endocrine cells arise from a generalized population and not from insulin-producing cells (Arntfield and van der Kooy 2011). On the other hand, the co-option of a program expressed by a β pancreatic-like cell is supported by the fact that several transcription factors expressed in pancreatic progenitors are also reutilized to guide the differentiation of the β-cells and that β-endocrine cells can proliferate and give

rise to new β-cells (Arntfield and van der Kooy 2011). However, the only way to validate the extent of that theory and the type of neuronal program co-opted would be to identify when specialized β-endocrine gut cells appeared during evolution and investigate the neuronal and non-neuronal expression profiles of pancreatic-like genes.

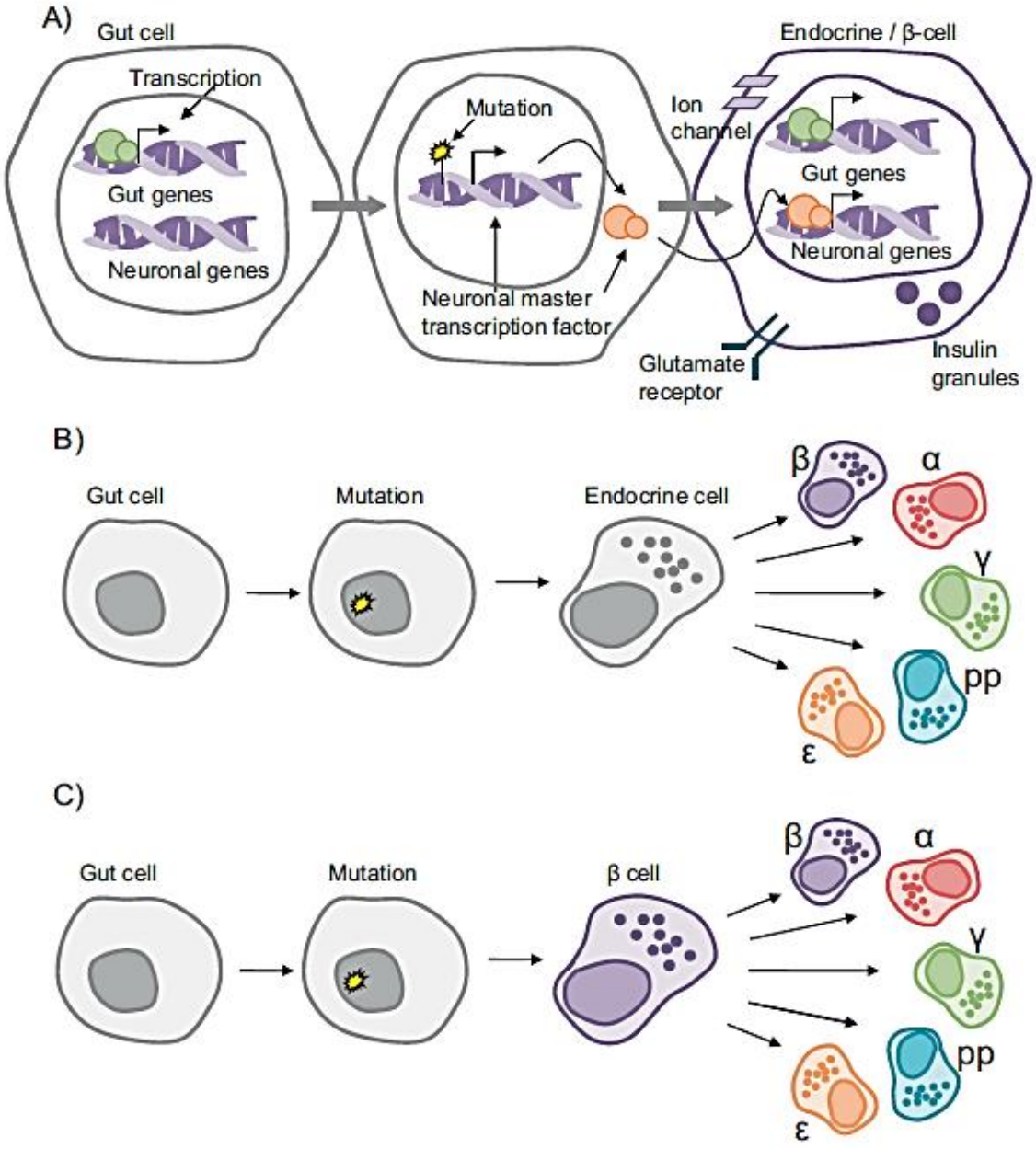


Figure 5.3. Proposed model describing the evolution of the endocrine pancreas (Arntfield and van der Kooy 2011). A) Convergent evolution model by which the mutation on the promoter of a neuronal master transcription factor regulator allowed its activation on gut cells. B) Model by which the initial endocrine cells adopted a common neuronal program giving rise to a common endocrine cell. C) Model by which the initial endocrine cell adopted a specialized neuronal regulatory program resembling a β-endocrine cell rather than a common endocrine cell.

5.4. Pancreatic-like cell types in sea urchins

Even though the formation of the pancreas as a distinct organ is a vertebrate feature, gene expression and functional studies identified several pancreatic-like cell types in animals outside this clade, including echinoderms.

In the sea urchin *S. purpuratus*, cells located in the upper part of the larval stomach (exocrine pancreas-like cells) were found to co-express pancreatic exocrine genes involved in food digestion, while their gene regulatory program depends on the Notch signaling pathway and the *Sp-Ptf1a* activation (Perillo et al. 2016). Furthermore, it was demonstrated that apart from the gene expression conservation, the sea urchin homolog of the vertebrate *Ptf1a* gene can substitute for its vertebrate homolog in activating downstream gene targets (Perillo et al. 2016). In addition to the exocrine pancreas-like cells, sea urchin larva also contains specialized gut cells that produce a structurally similar to the cephalochordate amphioxus insulin-like peptide (Perillo and Arnone 2014). Interestingly, the insulin positive cells were found in the intestinal region of the larva, which is known to be patterned and controlled by the sea urchin homolog of the vertebrate Pdx1 gene (Cole et al. 2009).

Based on the presence of at least two distinct pancreatic-like cell types in the sea urchin larva, next, I set out to investigate which other genes involved in the development of the vertebrate pancreas can be traced back to sea urchin cell types and whether I could find any evidence supporting the co-option of a neuronal genetic program by pancreatic β-cells. Since the presence of the first differentiated pancreatic-like cell type (exocrine pancreas-like cells) in the sea urchin larva has been reported at the 3 dpf *S. purpuratus* larva, my scRNA-seq analysis was focused on that specific time-point.

The first step to address those questions was to plot the average expression of sea urchin homologs of vertebrate transcription factors involved in distinct and representative steps of pancreas formation. Plotting for sea urchin homologs of the vertebrate transcription factors *FoxA2*, *Sox9, Gata6, Ngn3* expressed by pancreatic progenitors, the transcription factors *Pdx1, Nkx6.1, Pax6, NeuroD1, Mnx, Islet, Nkx2.2, Rfx3, Rfx6* guiding β-cell differentiation, the transcription factor *Brn4* promoting the production of glucagon in α-cells and genes exclusively expressed in acinar cells such as the transcription factors *Ptf1a*, *Rbpj, Mist,* as well as genes encoding form zymogenes *Cpa, Amy, Pnlip,* allowed me to assess their presence and putative enrichment in distinct cell types (Figs. 5.4, 5.5).

5.4.1. Exocrine pancreas-like cells

Ptf1a, Mist, Cpa, Amy, and *Pnlp* transcripts according to the scRNA-seq analysis are predicted to be all differentially expressed in a single putative broad cell type, the exocrine pancreas-like, as previously demonstrated by FISH in Perillo et al. 2016, validating that the single cell approach can be used as a means to identify pancreatic-like identities (Figs. 5.4, 5.5 H). In addition to the already known exocrine pancreas-like gene markers, the transcription factor *Rbpj*, which is a member of the *Ptf1a* complex, that binds to the promoter of other acinar specific genes as well as of the digestive enzymes was found to be expressed in the exocrine pancreas-like cell type (Fig. 5.4). The presence of *Rbpj* transcripts in the same putative broad cell type as *Ptf1a*, and the ability of the sea urchin homolog of the vertebrate *Ptf1a* gene to normally interact with the mammalian *Rbpj* (Perillo et al. 2016), implies that similar regulatory interactions between *Rbpj* and *Ptf1a* might take place in the sea urchin larva exocrine cells. Notably, expression of the transcription factors *Pax6* (Fig. 5.5B) and *Gata6*, involved in endocrine (Mitchell et al. 2017) and exocrine pancreas (Martinelli et al. 2013), respectively, was also detected in the exocrine pancreas cells (Fig. 5.4). It has been demonstrated by several studies that silencing of *Pax6* results in loss of β-cell due to the inability of cells constituting the islet of Langerhans to properly differentiate (Hart et al. 2013, Mitchell et al. 2017). Its presence in the exocrine-like pancreas cell type might reflect either an unknown role in exocrine cell differentiation or a reminiscent of an ancient role of *Pax6* was in promoting the differentiation of exocrine tissues that was evolutionary shifted towards promoting endocrine fate.

5.4.2. Endocrine pancreas-like cells

While the genes composing the molecular signature of the exocrine pancreas-like cell type were all found to be enriched in a single cell type, the rest of the transcription factors were broadly expressed in several cell types suggesting that they have a broader function in sea urchins than their chordate counterparts. Interestingly, a significant overlap of these transcription factors, especially those promoting β-cell differentiation, was found in the neuronal, intestinal, and pyloric sphincter putative broad cell types (Fig. 5.4). This is an interesting observation since it implies that β-endocrine-like genetic programs are operating in both the neurons

and gut cells of an invertebrate, making the sea urchin larva an ideal model to study whether the possible co-option of a genetic program was based on a pan-neuronal or specialized program.

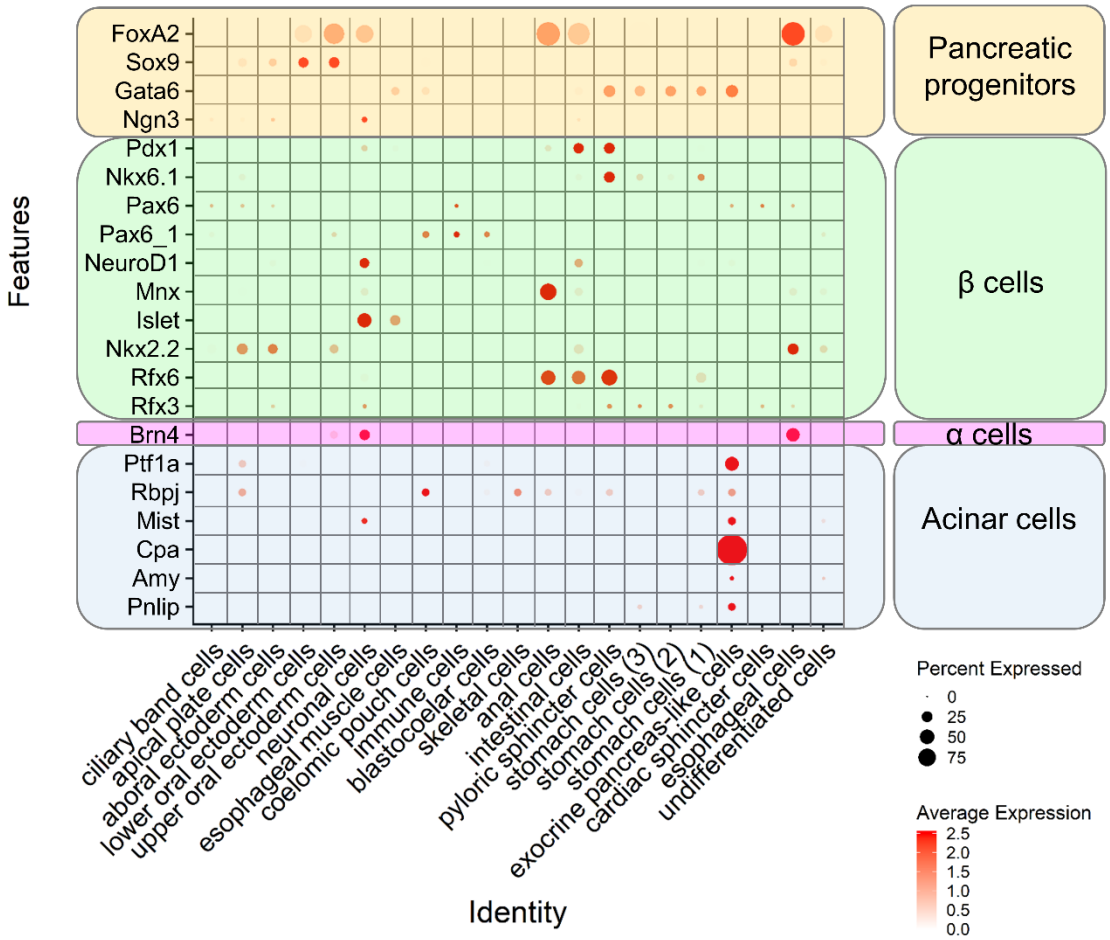


Figure 5.4. Pancreatic-like cell types in the sea urchin larva. Dotplot showing the average expression of a subset of genes whose chordates homologs are known markers of pancreatic endocrine progenitors, β-cells, α-cells, and acinar cells.

The combinatorial presence of *Pdx1*, *Nkx2.2, Nkx6.1, Mnx, NeuroD1, Rfx3,* and *Rfx6* transcripts in the two endodermally derived putative broad cell types, in which also insulin has been detected by FISH and IHC at later larval stages (Perillo and Arnone 2014), suggests that those genes are part of the gene regulatory network controlling the differentiation of the insulin-producing cells. Moreover, the transcription factors *FoxA* and *Gata6* involved in the initial steps of pancreas progenitors formation are found in the intestinal and pyloric sphincter cell types, respectively (Decker et al. 2006, Zinovyeva et al. 2016). On the other hand, no insulin transcripts were found in any of the single cell analyzed datasets, which is in line with previous reports demonstrating that insulin starts to be expressed from the

10 dpf larva stage and onwards (Perillo and Arnone 2014). Nonetheless, intestinal and pyloric sphincter putative broad cell types share both molecular and functional characteristics with the β-endocrine cells, including their dependence on the function of *Pdx1*, since silencing of *Pdx1* results in gut malformations. Such malformations include the loss of the pyloric sphincter, lower stomach capacity, and misdifferentiated intestinal cells (Cole et al. 2009, Annunziata et al. 2014).

Taking into account the expression pattern of these genes, as revealed by the scRNA-seq analysis, the role of their chordate homologs in the differentiation of β-cells (Sharma et al. 1999, Wendik et al. 2004, Doyle and Sussel 2007, Ait-Lounis et al. 2010, Smith et al. 2010, Perelis et al. 2015, Ray et al. 2016), and the late production of insulin in the same larval domains, lead to a possible scenario by which these two endodermal putative broad cell types with a β endocrine cell-like molecular signature are still differentiating. However, more information is needed to reach a safe conclusion on the differentiation status of these cells at this time-point and to identify which is the molecular signature of the cells that produce insulin later in development.

Apart from endodermally derived cell types exhibiting a strong β endocrine-like pancreatic signature, the pancreatic gene markers *FoxA*, *Ngn3, Pdx1, NeuroD1, Islet, Rfx3, Brn4,* and *Mist* were differentially expressed also in the neuronal putative broad cell type (Fig. 5.4). At the same time, neuronal expression of these genes has also been reported in several animal models (Blader et al. 2004, Ferri et al. 2007, Park et al. 2008, Sun et al. 2008, Boutin et al. 2010, Simon-Areces et al. 2010, Benadiba et al. 2012, D'Amico et al. 2013). Furthermore, the involvement of the sea urchin homologs of *Brn4* and *Ngn3, Brn1/2/4,* and *Ngn* has been already well established with *Brn1/2/4* promoting the differentiation of all neurons through a transient expression followed by the expression of terminal differentiation genes (Wei et al. 2016, McClay et al. 2018) and *Ngn* being necessary for the specification of cholinergic ciliary band neurons (Slota and McClay 2018). Moreover, FISH data have previously demonstrated that *Pdx1* is expressed apart from the endoderm (Cole et al. 2009) also in ectodermal cells alongside *Brn1/2/4* in *S. purpuratus* larva (Cole and Arnone 2009).

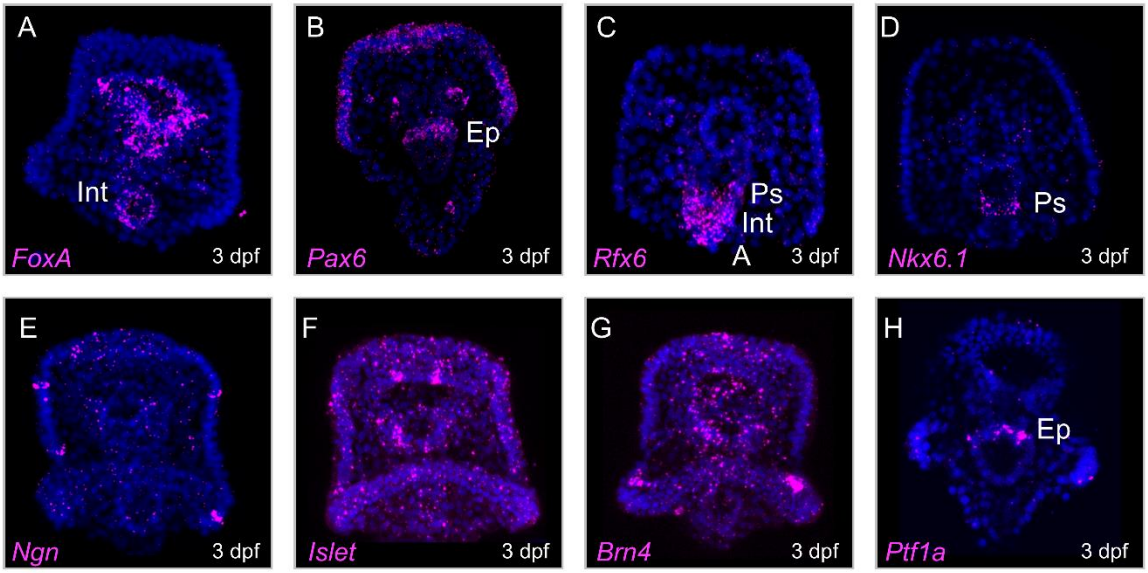


Figure 5.5. Expression patterns of pancreatic transcription factors. FISH using antisense RNA probes for *FoxA* (A), *Pax6* (B), *Rfx6* (C), *Nkx6.1* (D), *Ngn* (E), *Islet* (F), *Brn4* (G), and *Ptf1a* (H). Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. A, anus; Ep, Exocrine pancreas; Int, intestine; Ps, the pyloric sphincter

To identify the spatial distribution and possible co-expression patterns of the pancreatic transcription factors that were expressed in the neuronal putative broad cell type, I plotted for their average expression in the neuronal subclustering Seurat object (Fig. 5.6). This analysis revealed that although the pancreatic-like transcription factors were broadly expressed across the nervous system, three subpopulations were defined by the differential combinatorial expression of *FoxA2, Ngn3, Islet, Brn4 Pdx1, Brn4*, *Mnx,* and *Mist* (Fig. 5.6). Furthermore, the co-localization of *Pdx1* and *Brn4* in one neuronal population proposed that the ectodermal cells previously identified as *Pdx1/Brn4* double-positive ectodermal cells are indeed neurons.

According to the theory supporting the convergent evolution of β-endocrine cells and neurons, β-pancreatic cells arose either from the gut adopting a general neuronal program or a specific one (Arntfield and van der Kooy 2011). Subsequently, these gut cells would either bear characteristics resembling pluripotent endocrine cells giving rise to diverse endocrine types or specialized in making specific gene products. Based on the expression profile and the enrichment of the pancreatic transcription factors in three neuronal populations, it seems that three different pancreatic-like gene regulatory networks are present in the sea urchin nervous system. These neuronal populations contain transcription factors activated during distinct steps of the mammalian endocrine pancreas differentiation process,

although these genes have never been shown to be expressed in such a combinatorial way. Nonetheless, it is evident that an endocrine pancreas-like molecular signature is possibly guiding neuronal differentiation and function.

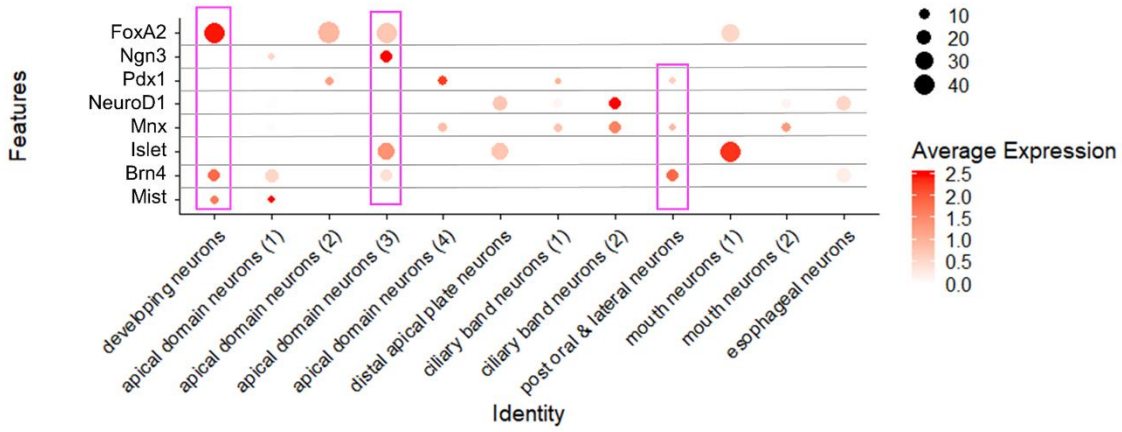


Figure 5.6. Average expression of pancreatic transcription in the nervous system of the 3 dpf pluteus larva. Dotplot showing the average expression of the transcription factors *FoxA2, Ngn3, Pdx1, NeuroD1, Mnx, Islet, Brn4, Mis*t, and their distribution in sea urchin larva neurons.

Based on the dominant role of the mammalian homolog of Sp-*Pdx1* during both the initial specification of pancreatic progenitors and during the differentiation and maturation of β-endocrine cells (Wilson et al. 2003, Fujimoto and Polonsky 2009, Zhu et al. 2017), its significance in promoting the differentiation of sea urchin gut endocrine-like cell types (Cole et al. 2009, Annunziata et al. 2014) and its expression in endodermal insulinergic cells during late larval development (Perillo 2013), the attention was focused in neuronal cell types expressing *Sp-Pdx1*.

5.5*. Sp-Pdx1* is expressed in differentiated neurons

As mentioned above, FISH data have demonstrated that *Sp-Pdx1* is expressed in the posterior endoderm and co-expressed with *Sp-Brn1/2/4,* the sea urchin homolog of the mammalian α-cells differentiation marker *Brn4*, in distinct ectodermal cells (Cole and Arnone 2009, Cole et al. 2009). Performing IHC experiments using an anti-Sp-Pdx1 antibody (Annunziata and Arnone 2014) and the marker of differentiated neurons Synaptotagmin B (Burke et al. 2006), showed that these cells are neurons (Figure 5.7 A-C), which is in line also with the scRNA-seq analysis performed on an earlier larval stage. Based on their transcriptomic profile (Figure

4.9) and spatial location, these neurons correspond to the larval lateral/post-oral neurons (Figure 5.7 A-C).

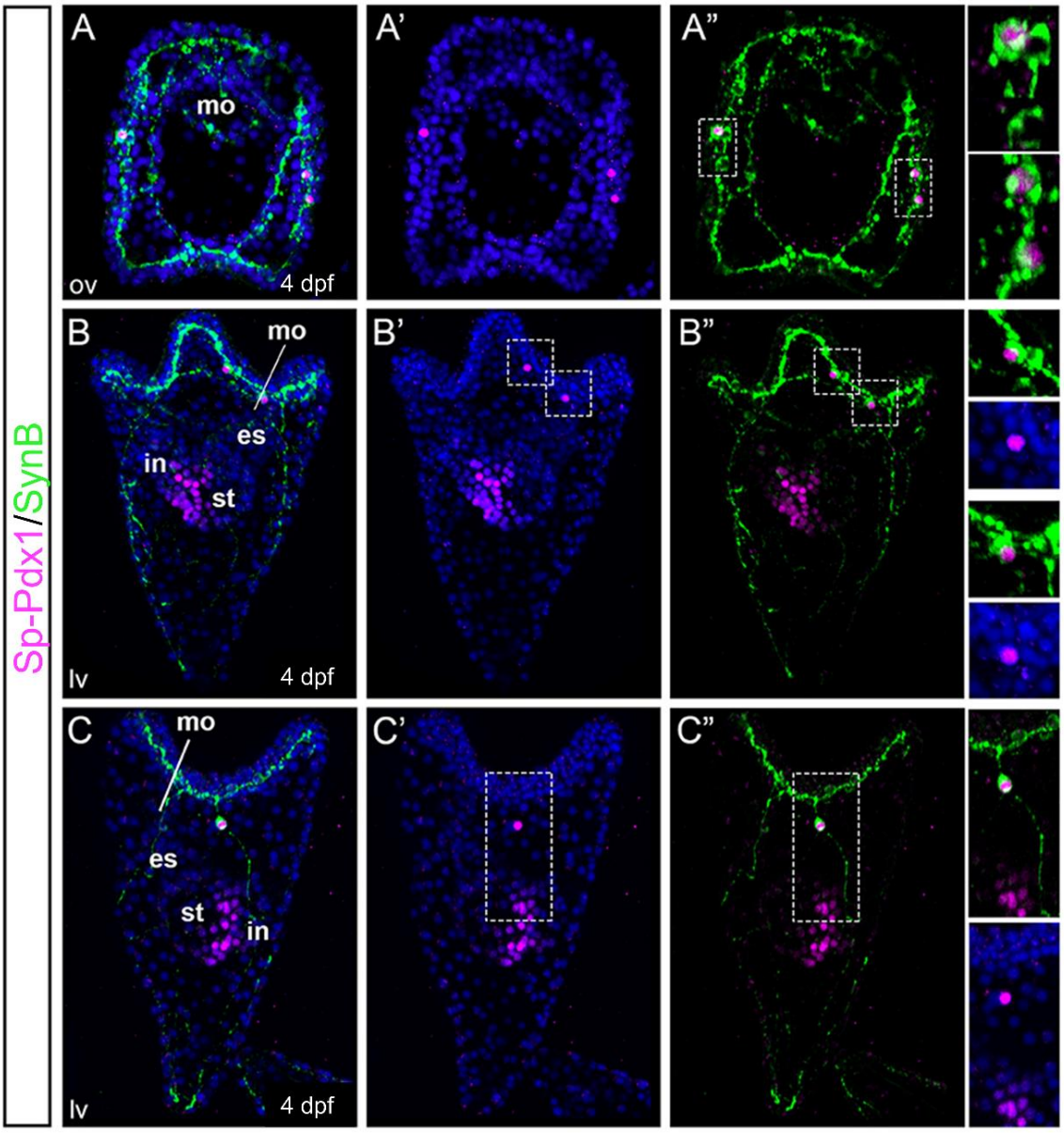


Figure 5.7. Neuronal expression of *Sp-Pdx1*. (A–C") Sp-Pdx1 (magenta) and SynB (green) protein localization in late larvae. Insets on the right highlight the expression of Sp-Pdx1 in the nuclei of the cells. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. es, esophagus; in, intestine; mo, mouth; st, stomach; lv, lateral view; ov, oral view (Figure adapted from Perillo et al. 2018).

While Sp-Pdx1 protein was detected only in the lateral neurons of the larva, FISH using an antisense probe for *Sp-Pdx1*, revealed a more dynamic expression in distinct cells of the foregut, the *Sp-Brn1/2/4* positive lateral and post oral neurons, as well as cells associated with the anterior/apical domain of the embryo and larva. At gastrula stage, *Sp-Pdx1* transcripts were also detected in cells within the anterior

neuroectoderm domain (Fig. 5.8 A), while at pluteus stage, expression in the apical domain was still present although restricted to a subset of cells (Fig. 5.8 B). Interestingly, scRNA-seq was able to detect *Sp-Pdx1* transcripts only in one population of cells constituting the anterior neuroectoderm at gastrula stage, while no transcripts were found in any apical plate populations (Fig. 5.9). To investigate whether the signal obtained by FISH in the apical domain of pluteus is an artifact of the technique or indication of real expression, I considered the possibility that *Sp-Pdx1* is expressed in the neuronal cells harbored within the apical plate. Indeed, plotting of *Sp-Pdx1* using the object containing the neuronal subclustering for the 3 dpf larva showed that *Sp-Pdx1* is found in two apical plate associated neuronal populations (Fig. 5.9).

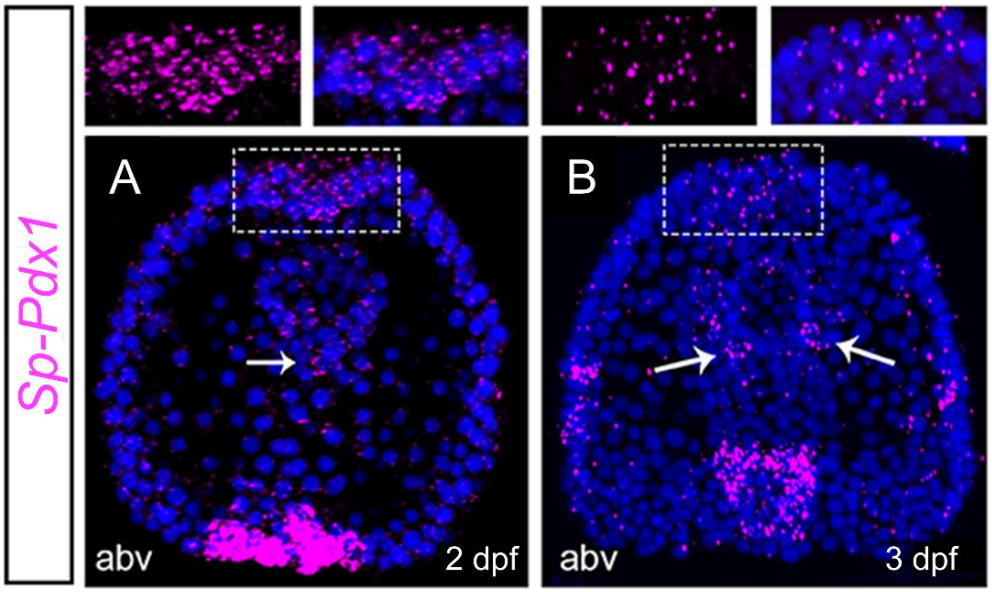


Figure 5.8. FISH localization of *Sp-Pdx1* transcripts. FISH using an antisense RNA probe for *Sp-Pdx1* at gastrula (A) and pluteus stages (B). Insets show the localization of *Sp-Pdx1* mRNA in the anterior neuroectoderm and apical plate of gastrula and pluteus stages. White arrowheads indicate the localization of *Sp-Pdx1* transcripts in the foregut. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. abv, aboral view (Figure adapted from Perillo et al. 2018).

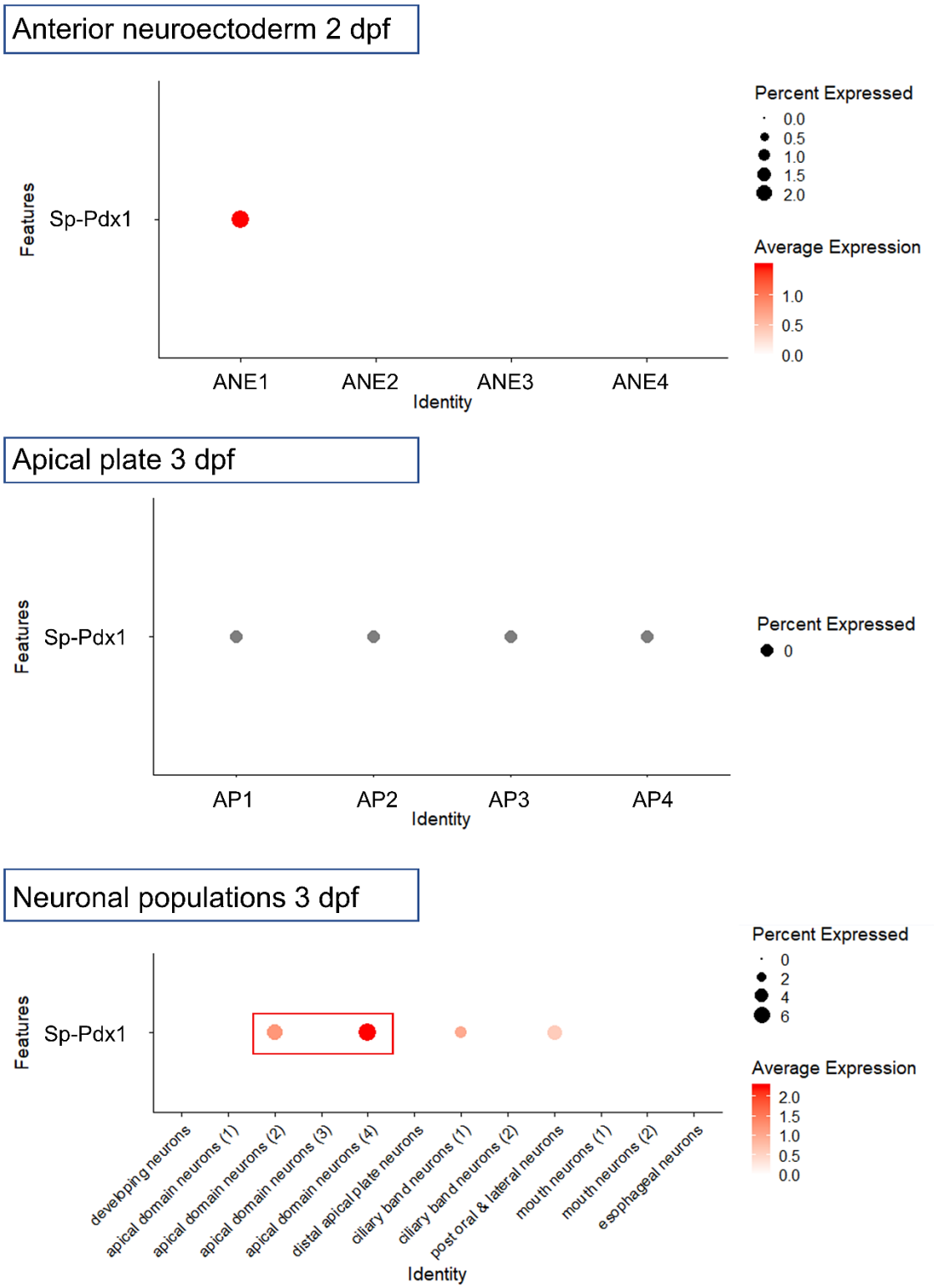


Figure 5.9. Expression of Sp-Pdx1 in the anterior neuroectoderm, apical plate, and apical plate neurons. Dotplot showing the average expression of Sp-Pdx1 in the anterior neuroectoderm, apical plate, and neuronal subclustered putative broad cell types.

The expression of *Sp-Pdx1* in the pre-neuronal anterior neuroectoderm domain at gastrula stage, combined with its expression only in differentiated neurons deriving from that region at the pluteus stage, implies that *Sp-Pdx1* might be acting in waves

similar to what has been described in mammals. In detail, *Pdx1* is amongst the first genes to be activated in the mammalian primitive gut tube promoting pancreas formation, then goes silent, and its expression is turned on during the differentiation of the β-endocrine cell lineage (Fujimoto and Polonsky 2009). Supposing that *Sp-Pdx1* functions in a similar way, as its transient expression in pre-neuronal and neuronal domains suggests, this would mean that *Sp-Pdx1* is necessary first for the proper specification of anterior neuroectoderm/apical plate domains and is reutilized again later for the differentiation of the neuronal types that arise from them.

To further investigate such a role, it is essential to identify the terminal differentiation genes produced by the *Sp-Pdx1* positive neurons and thus their identity. According to studies on sea urchin neuropeptides (Rowe and Elphick 2012, Wood et al. 2018), the expression domain of the neuropeptide *Sp-An* resembles the expression pattern of the *Sp-Pdx1/Sp-Brn1/2/4* double-positive neurons, while it has also been shown to be produced by some cells embedded within the apical organ: the serotonergic neurons. Indeed, scRNA-seq shows that *Sp-Pdx1* is co-expressed with *An* in both neuronal domains (Figure 4.9). In addition to the information that both *Sp-Pdx1* positive and *Sp-Pdx1/Sp-Brn1/2/4* double-positive neurons express the neuropeptide *Sp-An*, it was also predicted that the apical plate *Sp-Pdx1/Sp-An* neurons express *Sp-Tph*, suggesting they can produce the neurotransmitter serotonin as it was previously demonstrated through co-localization of *Sp-An* and serotonin in apical plate neurons (Wood et al. 2018). Regarding the *Sp-Pdx1/Sp-Brn1/2/4/Sp-An* triple-positive neurons *Sp-Th* and *Sp-Chat*, which are involved in the biosynthesis of dopamine and acetylcholine, respectively, were among the genes predicted to be expressed in this neuronal type. Notably, according to the scRNA-seq analysis, both *Sp-Pdx1* positive neuronal populations can produce and secrete complex and distinct neuromodulator cocktails.

Performing FISH for *Sp-An, Sp-Pdx1,* and *Sp-Brn1/2/4*, confirmed the scRNA-seq predictions and showed that *Sp-An* and *Sp-Pdx1* are co-localized both in the apical plate and post-oral/lateral neurons, while *Sp-Pdx1/Sp-Brn1/2/4/Sp-An* triple-positive cells corresponded to the post-oral/lateral ganglia neurons (Fig. 5.10).

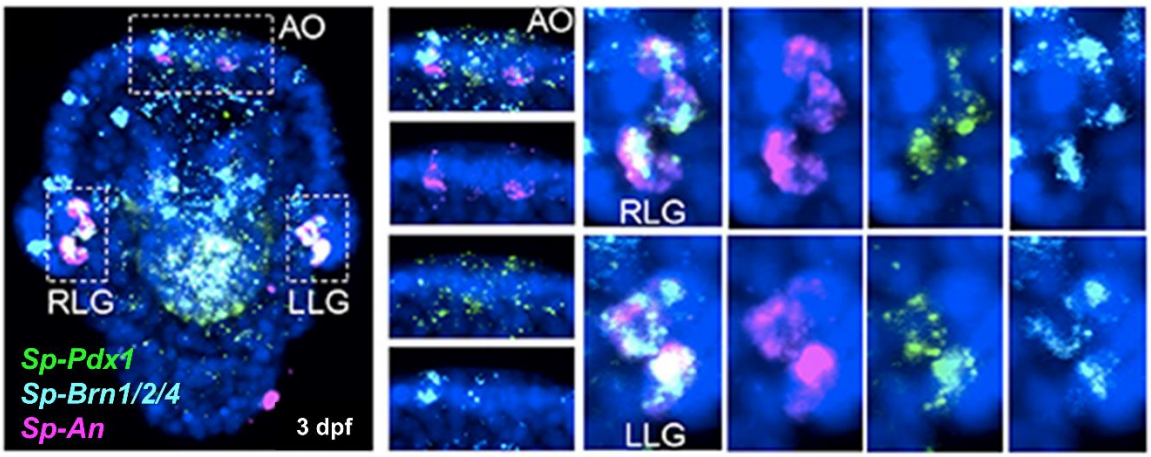


Figure 5.10. Co-localization analysis for *Sp-Pdx1, Sp-Brn1/2/4,* and *Sp-An* transcripts. Triple FISH showing the expression domains of *Sp-Pdx1* (green), *Sp-Brn1/2/4* (cyan), and *Sp-An* (magenta). Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. AO, apical organ; LLG, left lateral/post-oral neurons; RLG, right lateral/post-oral neurons. (Figure adapted from Perillo et al. 2018).

5.6. *Sp-Pdx1* is involved in neuronal differentiation

After identifying the neuronal characteristics of the *Sp-Pdx1* neurons, I set out to investigate whether *Sp-Pdx1* regulates the expression of the neuromodulators expressed by them. To this end, I used the already available differential transcriptomic analysis for *S. purpuratus Pdx1* morphants (Annunziata and Arnone 2014) that was performed at the same developmental stage as the scRNA-seq datasets (3 dpf). If *Sp-Pdx1* regulates the expression of Sp-An and Sp-Tph (thus serotonin), both of these genes should be downregulated compared to control larvae. Interestingly, according to the differential RNA-sequencing analysis, both genes were downregulated. Simultaneously, the absence of *Sp-An* and serotonin was also assessed by repeating the morpholino experiments and testing *in vivo* the localization of *Sp-An* transcripts and protein and serotonin mediated immunofluorescence (Figs. 5.11, 5.12). In *Sp-Pdx1* morphants, the number of *Sp-An* positive neurons was significantly reduced at 3 and 7 dpf. For instance, control larvae had an average of 3 *Sp-An* positive apical plate neurons, while *Sp-Pdx1* morphants had either zero or 1 apical plate neurons (Figure 5.11 A, A', B, B').

Similarly, the lateral/post-oral Sp-*Pdx1*/*Sp-Brn1/2/4* positive neurons had fewer Sp-An positive neurons. Also, in the case of serotonin, *Sp-Pdx1* morphants had fewer serotonergic neurons compared to control larvae. In detail, control larvae had an average of 4 bilaterally symmetric serotonergic cells in the apical organ, while Sp-

Pdx1 morphants had an average of two (Figure 5.11 C, C'). These results suggest that *Sp-Pdx1* regulates the expression of neuromodulators in the cells they are co-expressed.

Figure 5.11. *Sp-Pdx1* regulates the production of serotonin and *Sp-An*. Bar plot showing the downregulation of *Sp-An* and *Sp-Tph* as revealed by differential RNA sequencing analysis of *Sp-Pdx1* knockdown larvae. The data can be found in the non-book component (Sp-Pdx1 differential RNA-seq).

All the evidence gathered so far in terms of sea urchin neuronal specification implies that, by early gastrula stage, all neuronal precursors that will give rise to the nervous system of the early pluteus larva have been already specified. Taking into account that according to the *S. purpuratus* developmental bulk transcriptome database, *Sp-Pdx1* starts to be expressed at approximately 40 hpf (Tu et al. 2014), an early role in the specification of neuronal identity seems unlikely. Additionally, double FISH of *Sp-Pdx1* and *Sp-SoxC*, which is the first transcription factor known to be expressed by specified neurons, shows that *Sp-SoxC* is expressed in neuronal progenitors at 40 hpf, while *Sp-Pdx1* transcripts are absent at this stage and reach a significant level of expression at 48 hpf (Fig. 5.13). These data highlight that by the time *Sp-Pdx1* is turned on, the cells expressing it have already been specified as neurons. While neuronal fate is sealed at early gastrula (approximately 27 hpf), terminal differentiation genes indicative of a distinct neuronal function are found to be expressed from the late gastrula stage (48 hpf) and onwards (Burke et al. 2014, Garner et al. 2016, Wei et al. 2016, McClay et al. 2018). Based on this developmental window in which neuronal differentiation occurs and the effect of *Sp-*

Pdx1 knockdown as indicated in this work, a role of Sp-Pdx1 in neuronal differentiation is favored.

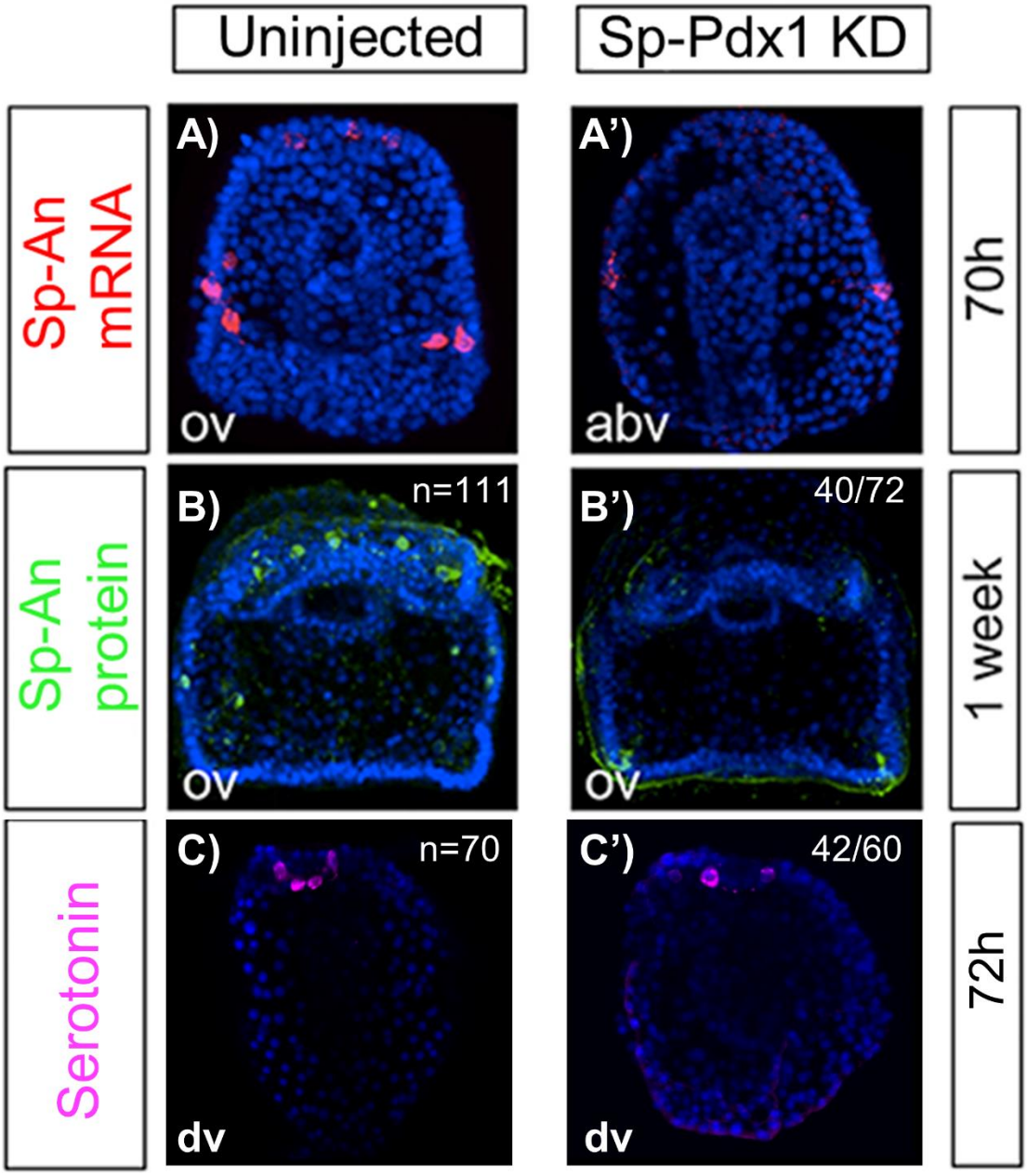


Figure 5.12. *Sp-Pdx1* is necessary for the production of *Sp-An* and serotonin. *Sp-An* mRNA detected by FISH in control (A) and *Pdx1* knockdown (A') larvae at 3 dpf. *Sp-An* protein detected by IHC in control (B) and *Pdx1* knockdown (B') larvae at 1 wpf. Immunohistochemical detection of serotonin in control (C) and *Pdx1* knockdown (C') larvae at 3 wpf. Numbers on the top right corner indicate the number of larvae analyzed and the amount of larvae with similar staining patterns. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. abv, aboral view; dv, dorsal view; ov; oral view. (Figure adapted from Perillo et al. 2018).

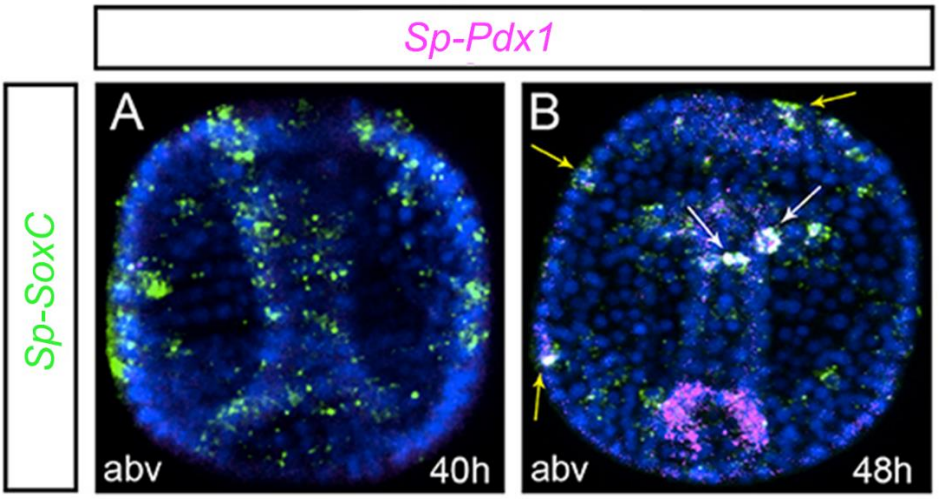


Figure 5.13. *Sp-Pdx1* positive cells are first specified as neurons. Double FISH of *Sp-Pdx1* and the proneural transcription factor *Sp-SoxC* in 40 (A) and 48 (B) hpf gastrulae. White arrows indicate *Sp-Pdx1* co-localization with *Sp-SoxC* in the foregut, yellow arrows show *Sp-Pdx1*, and *Sp-SoxC* double-positive ectodermally derived neural precursors. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. abv, aboral view. (Figure adapted from Perillo et al. 2018).

5.7. *Sp-Pdx1* as a regulator of neurosecretory fate

The scRNA-seq and the perturbation analyses presented in this Chapter of the thesis have indicated the presence of a pancreatic signature in the nervous system of the sea urchin larva. In particular, the transcription factor, *Sp-Pdx1,* a master gene controlling mammalian pancreas development and differentiation, is essential for proper differentiation of endodermal cell types (Cole et al. 2009, Annunziata et al. 2014). As revealed by this thesis's results, *Sp-Pdx1* is also required for the differentiation of distinct neuronal populations. In both cases of serotonergic and *Sp-An* positive neurons, knocking down of *Sp-Pdx1* impaired their neurosecretory ability.

To further investigate the role of *Sp-Pdx1* in the differentiation of the neuronal types it is expressed, I focused my attention on characterizing the *Sp-Pdx1/Sp-Brn1/2/4/Sp-An* triple-positive at a morphological, molecular, and functional level. The reasoning for this choice was based on the co-expression of *Sp-Pdx1* and *Sp-Brn1/2/4* in these cells and the essential role of *Sp-Pdx1* in regulating the expression of the Sp-An neuropeptide. Interestingly, the simultaneous expression of the transcription factors *Sp-Pdx1* and *Sp-Brn1/2/4* that in mammals orchestrate the differentiation of β and α cells from endocrine progenitors, respectively (Hussain et al. 2002, Wilson et al. 2003, Fujimoto and Polonsky 2009, Zhu et al. 2017), suggests

that these neurons could reflect a step in the endocrine cells evolution, in which only one multifunctional type of endocrine cells might have been present. For simplicity, *Sp-Pdx1/Sp-Brn1/2/4/Sp-An* triple-positive cells are hereby referred to as PPE neurons, which stands for “Pre-Pancreatic Endocrine” neurons, since they express genes whose homologs in mammals regulate the differentiation of α and β endocrine cells. According to their spatial position PPE neurons are referred to as Post Oral Neurons (PON) or Lateral Ganglion Neurons (LGN). PON are in close proximity to the post-oral arms of the larva, while LGN are localized on the left and right sides of the larva.

PPE neurons are bilaterally symmetric, with their cell bodies spatially placed towards the posterior end of the larva close to the post-oral arms region. Both IHC detection of the *Sp-An* protein and scanning electron microscopy of sliced larvae verified their spatial organization and the presence of axonal projections from these neurons towards the apical plate and ciliary band of the larva (Fig. 5.14). Moreover, double immunohistochemical staining of the differentiated neurons marker Synaptotagmin B and the Sp-An neuropeptide shows that the PPE neurons are in the vicinity of the ciliary band projecting axons towards the apical plate and the ciliary band itself (Fig. 5.15 A, B, C; see also Fig. 5.7). Double fluorescent *in situ* hybridization of *Sp-An* and *Sp-FbsL_2* as well as double immunostainings for Sp-An and acetylated tubulin, the former marking the ciliary band the latter labeling cilia, highlight the spatial distribution of these neurons relative to the ciliary band (Fig. 5.15 D, E, F). FISH for markers labeling the two different ciliary band neuronal subtypes as identified by this work, paired with immunostaining for the Sp-An protein, showed that not only the PPE neurons are projecting towards ciliary band cells but also to ciliary band neurons (Fig. 5.15 G, H, I). For instance, fluorescent *in situ* hybridization of *Sp-Prox1,* combined with immunostaining for Sp-An, clearly demonstrates that the axons projected from the PPE neurons are in contact with the cell bodies of the *Sp-Prox1* expressing neurons (Fig. 5.15 I).

The spatial organization of the PPE neurons resembles the domain in which the primary mesenchyme cells (PMCs) constituting the postoral skeletal rod are localized (Sun and Ettensohn 2014). In addition to the PMCs, globular cells, a specialized cell population involved in immune response, are also present in the same domain (Ho et al. 2017). To understand whether indeed PPE neurons are spatially related to these cell types, double IHC and FISH experiments were carried out. IHC staining of the skeletal cells marker Msp130 (Harkey et al. 1992) and the

neuropeptide An shows that the An positive neurons, closer to the post-oral arms, are close to the ventral-lateral cluster of PMCs (Fig. 5.15 J). Similarly, FISH *of Sp-MacpfA2*, which labels the globular cells (Ho et al. 2017), followed by immunohistochemical detection of the Sp-An protein, implies that globular cells are in contact with the post-oral An positive neurons (Fig. 5.15 K).

The transcription factors *Sp-Nk1* and *Sp-Otp,* as well as the catecholaminergic and cholinergic neuronal markers *Sp-Th* and *Sp-Chat* (Slota and McClay 2018), are among the genes predicted to be expressed in the PPE neurons, as revealed by the subclustering analysis presented in chapter 4. These predictions were verified by performing double FISH of Sp-*Nk1*, the neuropeptide Sp-An, and FISH of *Sp-Otp* paired with the immunostaining for Sp-An (Fig. 5.15 M, N).

Likewise, double immunostainings of the anti-Sp-An with anti-Th and anti-Chat antibodies showed co-localization of all three proteins in the PPE neurons (Fig. 5.15 L, O).

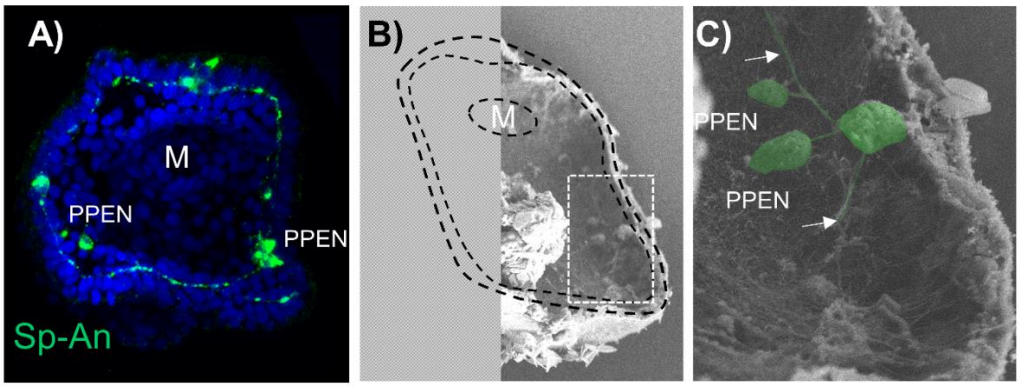
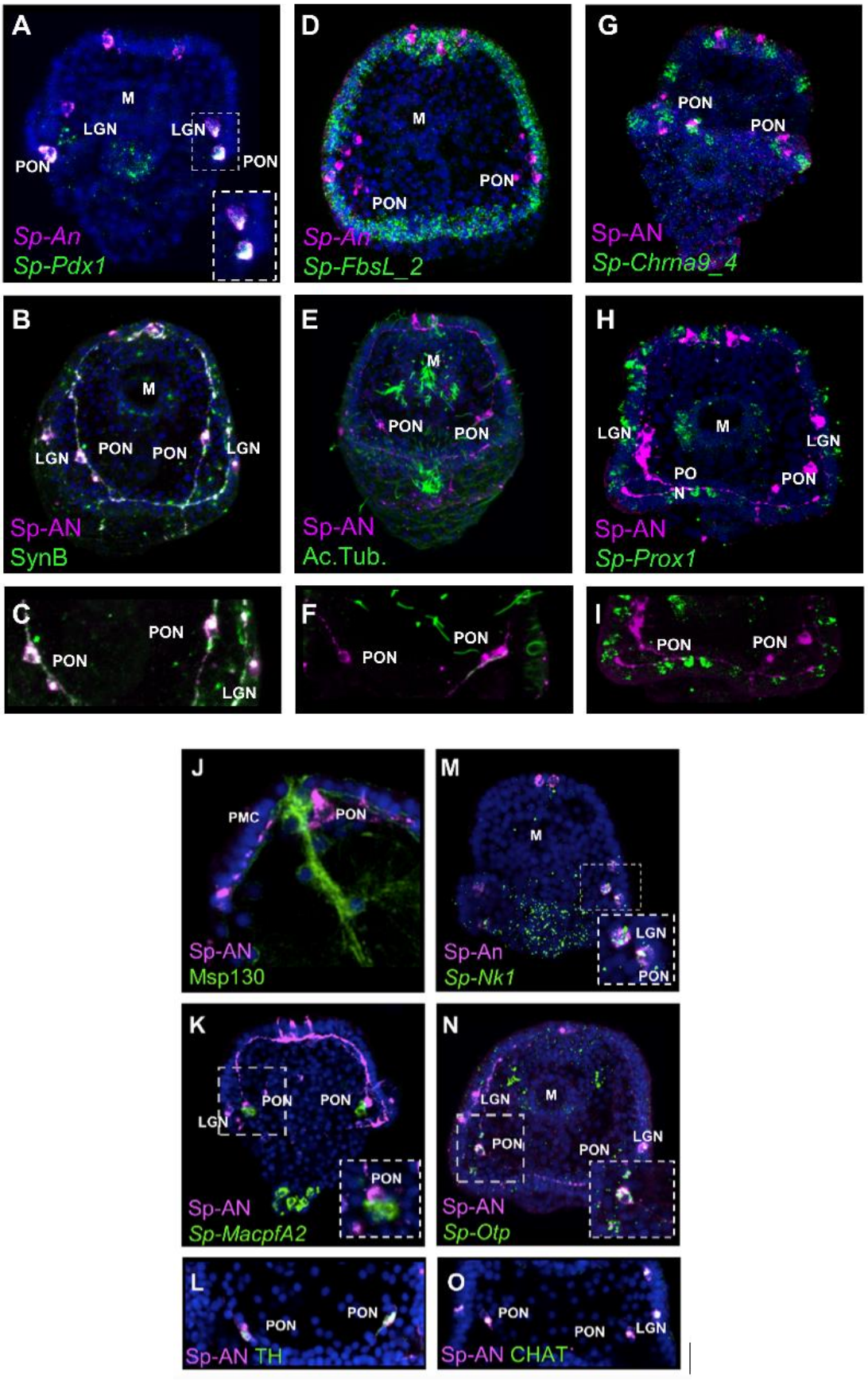


Figure 5.14. PPE neurons spatial localization in 3 dpf larvae. A) Immunohistochemical detection of the Sp-An neuropeptide. B) Scanning electron image of a sliced larva (X1500). Dotted lines indicate the shape of the larva, and the dotted circle indicates the mouth region. C) Higher magnification (X3000) of the same larva shown in B in which the PPEN neurons have been highlighted with pseudocolor (green). Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. Arrows indicate axonal projections. M, mouth; PPEN, pre pancreatic endocrine neurons.

Figure 5.15. Molecular characterization of PPE neurons at 3 dpf pluteus larva. A) Double FISH of *S. purpuratus* 3 dpf larvae for *Sp-Pdx1* and *Sp-An*. B) Immunohistochemical detection of the Sp-An and Synaptotagmin proteins. C) Close-up caption of the Sp-An PON neurons shown in B. D) Double FISH for *Sp-FbsL_2* and *Sp-An*. E) immunohistochemical detection of the Sp-An and acetylated tubulin proteins. F) Close-up caption of the Sp-An PON neurons shown in E. G) FISH of *S. purpuratus* 3 dpf larvae for *Sp-Chrna9_4* paired with IHC detection of Sp-An. H) FISH of *S. purpuratus* 3 dpf larvae with a specific antisense probe for *Sp-Prox1* combined with Sp-An immunostaining. I) Close up caption of the Sp-An PON neurons shown in H). J) Double immunostaining for the neuropeptide Sp-An and the skeletal cells marker Msp130. K) FISH for *Sp-MacpfA2* paired with immunohistochemical detection of Sp-An. L) Immunohistochemical staining for the neuropeptide Sp-An and the enzyme involved in dopamine biosynthesis Sp-Th. M) Double FISH for the transcription factor *Sp-Nk1* and *Sp-An*. N) FISH for the transcription factor *Sp-Otp* combined with immunohistochemical detection of Sp-An. O) Immunohistochemical detection of the

neuropeptide Sp-An and the enzyme Sp-Chat. Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. LGN, lateral ganglion neurons; M, Mouth; PON, Post-oral neurons. Note: the term PPE neurons reflects the molecular signature of this neuronal type, while the terms PON and LGN reflect their spatial location on the larva.



To further characterize the PPE neurons, I interrogated the scRNA-seq data looking for genes known to be involved in pancreatic development, particularly in β cell differentiation and function. This analysis identified 20 differentially expressed genes in the PPE neurons genes, all having a pivotal role in the formation and proper function of the mammalian endocrine pancreas (Fig. 5.16 A). The mammalian genetic program controlling the endocrine pancreas differentiation has been described in great detail, and many key transcription factors are essential for this process (Zaret and Grompe 2008, Tritschler et al. 2017).

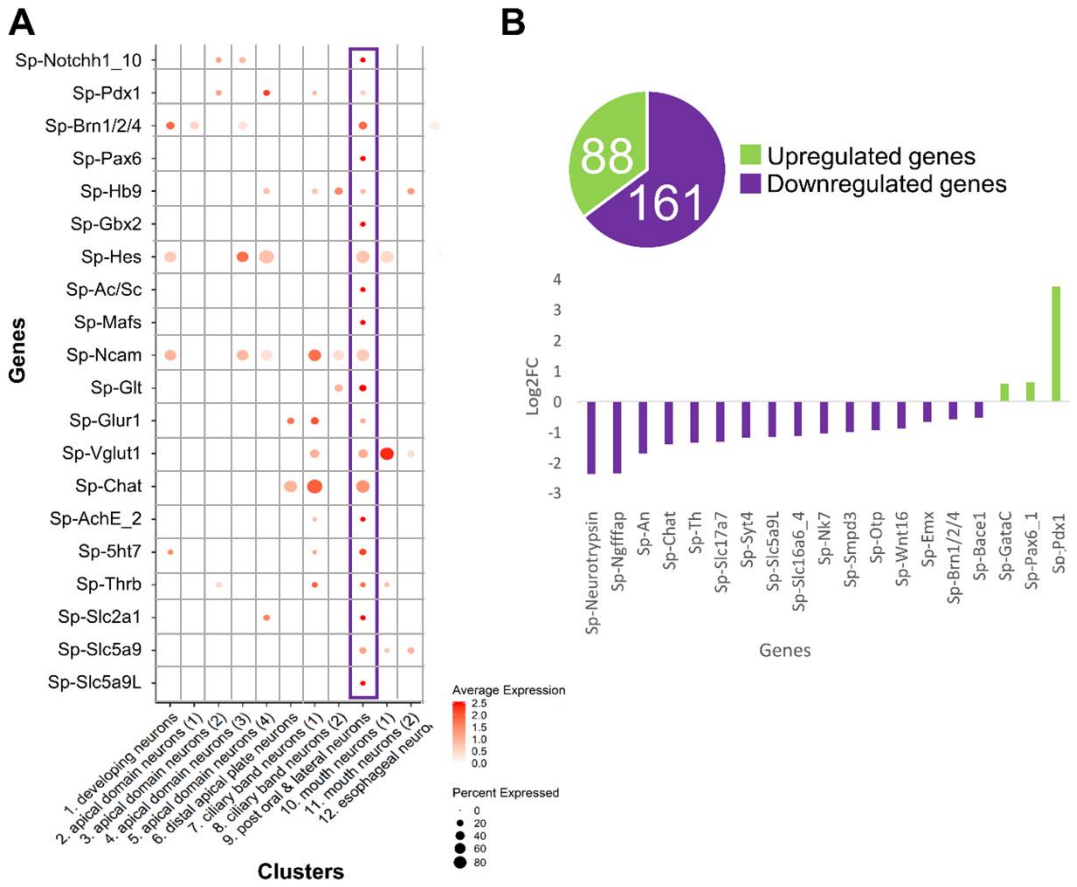


Figure 5.16. Molecular fingerprint of PPE neurons. A) Dotplot showing the enrichment of sea urchin homologs of vertebrate genes regulating endocrine pancreas formation in the PPE neurons. B) Bar plot of selected *Sp-Pdx1* target genes in the PPE neurons as revealed by differential RNA sequencing analysis of *Sp-Pdx1* morphants. The differential RNA-seq data come from Annunziata and Arnone 2014 and Lowe et al. 2016)

As previously mentioned, the initial steps of the pancreatic specification program include several signaling pathways, including Notch, which displays an important role in the determination of whether pancreatic progenitors will remain dormant or will differentiate towards a specific endocrine fate. Further steps of the differentiation process include the transcriptional activation of *Hes/hairy* and enhancer of split and

Hes1 that repress *Ascl1* (Ishibashi et al. 1995, de la Pompa et al. 1997, Iso et al. 2003). In particular, the transcription factor *Pax6* is high in the regulatory hierarchy establishing and maintaining the differentiated state of the mature β-cell (Hart et al. 2013, Mitchell et al. 2017, Buckle et al. 2018). In chordates, the silencing of the homeobox transcription factor Hb9 blocks the secretion of insulin, while the loss of its mice homolog results in abolishing the pancreatic differentiation program and partial agenesis (Li et al. 1999, Arkhipova et al. 2012). On the other hand, the transcription factor *Mafs* has been found to regulate insulin expression by competing for similar binding sites as insulin's activator *MafA*. Lastly, the transcription factor Gbx2 has been reported to be expressed in insulin-producing MIN6 cell lines (Mizusawa et al. 2004), in which its role remains unknown.

According to my analysis, the sea urchin homologs of Notch and the transcription factors mentioned above are transcribed in the PPE neurons. This implies that specific neurons of a non-chordate deuterostome follow a genetic program similar to the one utilized by endocrine pancreas cells and strengthens the hypothesis supporting the convergent evolution of β-pancreatic cells by co-optioning of a neuronal gene regulatory network by gut cells (Arntfield and van der Kooy 2011).

Additionally, transcripts for terminal differentiation genes, whose homologs are expressed both by the endocrine and nervous systems, were found in PPE neurons. Examples of such genes are the neural cell adhesion molecule *Ncam* known to be produced in both endocrine and nervous systems of the rat (Langley et al. 1989), as well as genes encoding components of the glutamate signaling pathway (*Sp-Glt*: glutamate synthase; *Sp-Vglut1*: glutamate transporter; *Sp-Glur1*: glutamate receptor), which in mammals is participating in the regulation of glucose-dependent insulin secretion (Gonoi et al. 1994, Maechler and Wollheim 1999).

Moreover, the single cell transcriptomics and molecular analyses of the PPE neurons identified genes involved in the neurotransmitter pathways of dopamine and acetylcholine (*Sp-Th*, *Sp-Chat,* and *Sp-AchE)*. The importance of their transcription in the PPE neurons comes from studies reporting that Th is present in the endocrine pancreas of multiple species (Iturriza and Thibault 1993, Teitelman et al. 1993), while choline acetyltransferase (*Chat*) has been found in human pancreatic islets. In the case of the produced acetylcholine, a specific role in the stimulation of insulin secretion by neighboring β-cells has been proposed (Rodriguez-Diaz et al. 2011).

Although no genes involved in the biosynthesis of serotonin were found in the PPE neurons, one serotonin receptor was found differentially expressed. This is an exciting finding since serotonergic signaling is thought to participate in the regulation of insulin secretion based on the observation that several serotonin receptors have been found to be expressed in human pancreatic islets (Amisten et al. 2015). Thus, serotonin may regulate the neuromodulators' secretion in a similar model as proposed for insulin release. Moreover, transcripts of the *Thyroid hormone receptor B* were found in the PPE neurons suggesting an interplay of thyroid hormone signaling. Interestingly, thyroid hormone signaling promotes endocrine pancreas differentiation in mice (Aiello et al. 2014), suggesting that thyroid hormone signaling might regulate the differentiation of PPE similarly.

The primary role of the endocrine pancreas cells is hormonal secretion, triggered by changes in extracellular glucose levels. Endocrine pancreas cells are equipped with Glucose transporter and co-transporters used to detect such changes (Navale and Paranjape 2016, Berger and Zdzieblo 2020). According to the scRNA-seq data, the PPE neuronal type contains transcripts for three glucose co-transporters, suggesting they have the machinery needed to detect changes in glucose levels similarly to endocrine pancreas cells.

Having an overview of the pancreatic molecular signature active in the PPE neurons, I set out to build a provision gene regulatory network that is active in this neuronal population and is focused on the role of *Sp-Pdx1*. To this end, I used the genes predicted to be expressed in the PPE neurons by the scRNA-seq analysis as a way to filter out the bulk *Sp-Pdx1* differential RNA-seq produced at the same developmental stage (Annunziata and Arnone, 2014, Lowe et al. 2016). In doing so, I found a total of 249 PPE genes, differentially expressed in the Sp-Pdx1 knockdown larvae, with 65% of the targets to be downregulated (Fig. 5.16 B). The transcription factors Sp-*Brn1/2/4* and *Sp-Otp*, both involved in sea urchin neuronal differentiation (Garner et al. 2016, Slota and McClay 2018), were downregulated, suggesting that their differentiation is compromised. In agreement with this, genes encoding either neuropeptides or enzymes involved in neurotransmitter biosynthesis such as *Sp-An, Sp-Ngfffap, Sp-Th*, and *Sp-Chat* were also downregulated, highlighting a severe functional impairment and loss of identity of the PPE neurons.

Based on this analysis, I drafted a provisional GRN depicting selected gene interactions and showing that *Sp-Pdx1* is essential for the differentiation of this cell

type and thus the acquisition of neurosecretory fate (Fig. 5.17). This GRN demonstrates the conserved role of *Sp-Pdx1* in promoting cell differentiation and establishing the secretory (in this case, neurosecretory) cellular identity. However, future studies are needed to verify these gene interactions, the hierarchy, and the actual connectivity of the GRN. Nonetheless, this GRN proves that combining scRNA-seq with various omics approaches effectively discovers potential gene interactions and regulatory wirings.

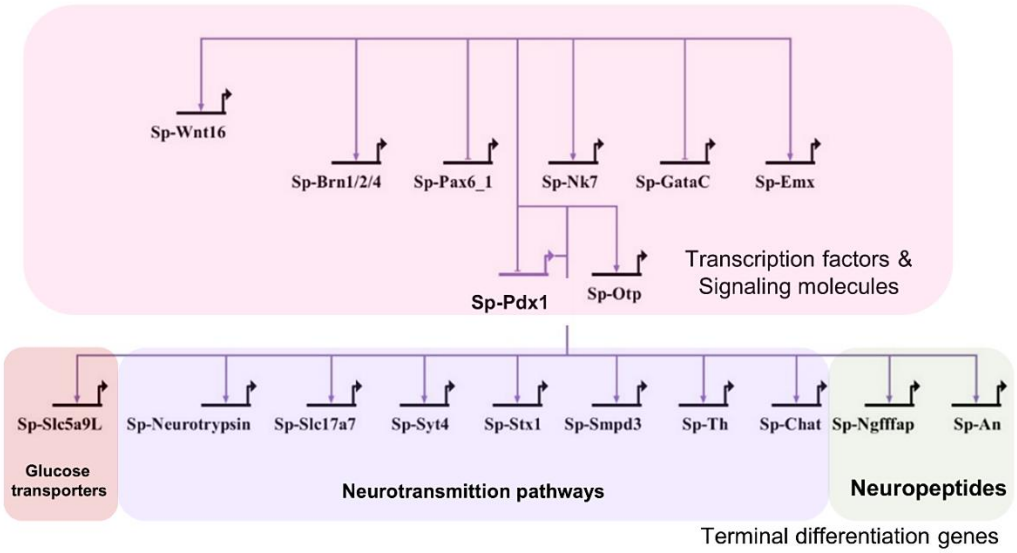


Figure 5.17. PPE neurons GRN. Provisional GRN operating in the Sp-PPE neurons as revealed by the combination of scRNA and differential RNA-seq analysis.

Altogether, the data presented in this section of the thesis show that key genes involved in the differentiation and function of the mammalian endocrine pancreas can be traced back in the PPE neurons. The presence of these genes, in a neuronal population of a non-chordate deuterostome that lacks the pancreas as a distinct organ, reinforces the hypothesis that pancreatic cells arose by recycling a pre-existing neuronal GRN. Furthermore, this program may resemble the one recognized in the PPE neurons and was most likely operating in the deuterostome ancestor.

However, what is this GRN used for in sea urchin? Are the PPE neurons also able to produce insulin? Based on the role of the mammalian homologs of the pancreatic genes found to be present in the PPE neurons, it seems that PPE neurons have the necessary machinery to produce insulin. As previously mentioned, no insulin transcripts could be traced back to our datasets being in line with the data presented

in Perillo and Arnone, 2014, showing that insulin transcripts are localized in the gut starting from 10 dpf and onwards. One possible scenario is that these neurons are still going through differentiation and will produce insulin later in development, which is in line with the late production of insulin in the gut that requires a 10-day differentiation window before insulin can be detected in the larval gut (Perillo and Arnone 2014). Another hypothesis is that the PPE neurons cannot produce insulin and that the machinery regulating insulin's production and secretion is reused for different functions. One of them could be controlling the neurosecretion of different hormones and neuromodulators. Based on the data presented in this work showing that *Sp-Pdx1* is essential for neuronal differentiation and that PPE neurons produce multiple neuromodulators (Fig. 5.16 A), the scenario in which this machinery is used to promote the differentiation of these neurons and thus their activity and production of neuromodulators (Sp-An, Sp-Ngfffap, Dopamine, Acetylcholine) is favored.

5.8. Investigating the role of the Sp-An neuropeptide

The overall analysis demonstrates that *Sp-Pdx1* has a major role in orchestrating neuronal differentiation in the neuronal types in which it is expressed, such as the PPE neurons (Figs. 5.16, 5.17) and a population of apical organ serotonergic neurons (Fig. 5.12). Notably, a common element expressed in both cell types and its production is equally impaired in *Sp-Pdx1* morphants is the neuropeptide *Sp-An* (Figs. 5.12, 5.16), suggesting that *Sp-Pdx1* is necessary for its production in two neuronal types with diverse regulatory landscapes. A way to investigate the importance of the *Sp-An* dependence on *Sp-Pdx1* expression in diverse neuronal types is to understand the role of the *Sp-An* neuropeptide. The only report on the role and evolution of the *Sp-An* neuropeptide comes from a study by Rowe and Elphick. According to the findings of this study, *Sp-An* is a 441-residue protein that consists of a 27-residue signal peptide and 13 copies of structurally related putative neuropeptides containing the N-terminal sequence Ala-Asn (AN). Interestingly, phylogenetic analyses could not indicate any similarities between *Sp-An* and neuropeptides from other phyla, suggesting that it is a neuropeptide specific to the echinoderm lineage (Rowe and Elphick 2012, Perillo et al. 2018, Wood et al. 2018, Carter et al. 2021).

To investigate the function of the Sp-An neuropeptide, I designed a specific Sp-An morpholino antisense oligonucleotide (MASO) that was injected into fertilized sea urchin eggs. The efficiency and the phenotype were assessed at 3 dpf pluteus larva, a stage in which the *Sp-Pdx1*/*Sp-An* co-localization and regulation was demonstrated (Fig. 5.18). This last section of the thesis contains preliminary results of these experiments indicative of the possible role of the *Sp-An* neuropeptide.

Immunohistochemical detection of the *Sp-An* protein in control (Fig. 5.18 B, C) and *Sp-An* morphants (Fig. 5.18 B', C') proved the successful silencing of *Sp-An* since no Sp-An mediated fluorescence was detected in the *Sp-An* MASO injected larvae. Moreover, light microscopy imaging of the *Sp-An* morphants did not show any severe morphological defects caused by the *Sp-An* knockdown in respect to control larvae (Fig. 5.18 A), apart from a strange morphology of the larval vertex (Fig. 5.18 A'). Confocal imaging of immunostained for beta-tubulin control and knocked down specimens led to similar conclusions (Fig. 5.18 B, B'). These morphological differences in the vertex structure (Fig. 5.18-yellow arrowheads) could be indicative of a cross-talk between An positive neurons either with the aboral ectoderm cells or the PMCs present regulating skeletal growth. In fact, as shown in Fig. 5.18 B (white arrows), neuronal projections containing Sp-An vesicles are found in the aboral ectoderm domain of the larva, supporting an An mediated control of skeletal/ectodermal growth and thus proper vertex formation. Nonetheless, it seems that the expression of the neuropeptide Sp-An is not crucial for morphogenesis and embryonic/larval development.

Next, I set out to investigate whether the knockdown of this neuropeptide affected the nervous system organization. Surprisingly, immunohistochemical staining for the differentiated neurons marker Synaptotagmin B showed less neuronal complexity in the Sp-An morphants in respect to control larvae (Fig. 5.18 C, C'). In detail, the number of Synaptotagmin B positive cells was reduced both in the apical organ and lateral domains in which An is expressed as well as the ciliary and mouth regions. Sp-An knockdown effects on the nervous system formation were also assessed by immunohistochemical detection of serotonin and the enzyme tyrosine hydroxylase (Th), typically co-produced with An in the apical organ and post-oral lateral neurons, respectively. This set of experiments indicated a reduction of serotonin positive neurons and complete loss of *Th* expression in the *Sp-An* morphants. Regarding serotonin, an average of three clustered serotonin positive

cells was found in the apical organ of the *Sp-An* morphants, while the apical organ of control larvae consists of 2 bilaterally symmetric pairs of serotonergic neurons.

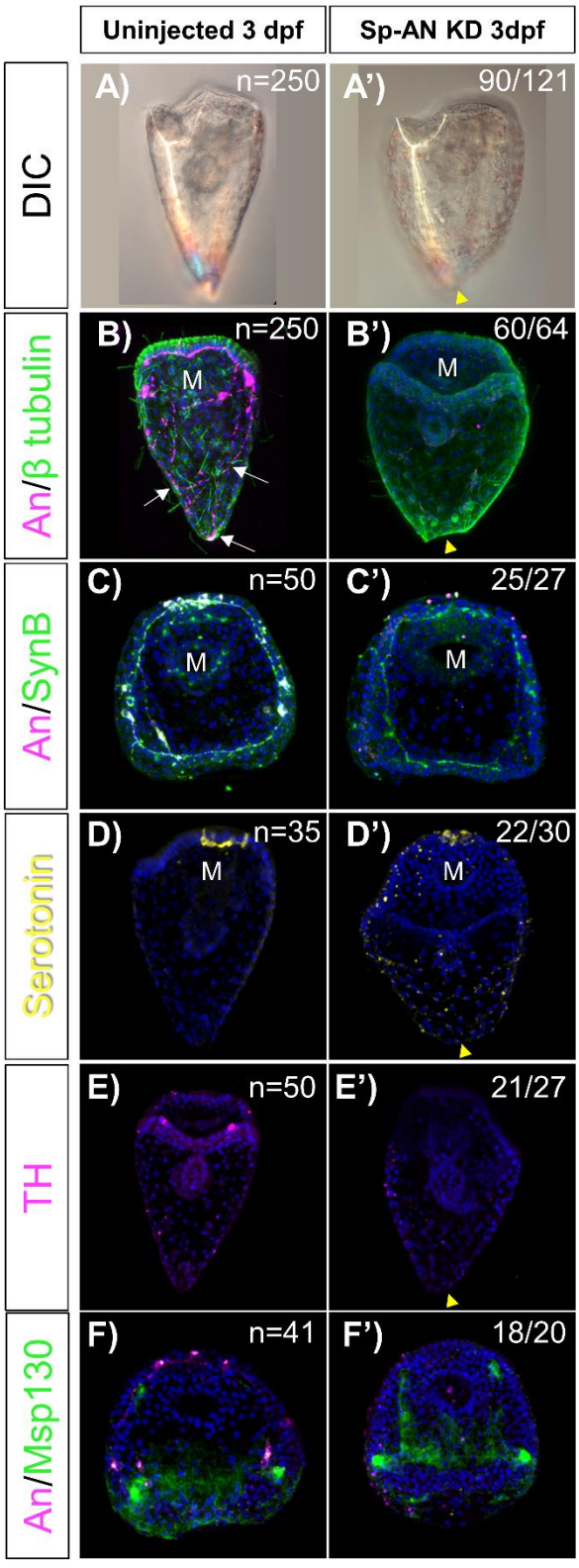


Figure 5.18. Morphological and molecular characterization of Sp-An knockdown larvae. (A-A') DIC images of 3 dpf control (A) and Sp-*An* knockdown larvae (A'). (B-B') Immunohistochemical detection of *Sp-An* (magenta) and *beta-tubulin* (green) in control (B) and Sp-An morphants (B'). (C-C') Immuno-histochemical detection of *Sp-An* (magenta) and *Synaptotagmin B* (green) in control (C) and Sp-An morphants (C'). (D-D') Immuno-staining for of serotonin in control (D) and Sp-An morphants (D'). (E-E') Immunohistochemical detection of *Th* in control (E) and Sp-An morphants(E'). Nuclei are labeled with DAPI (in blue). All images are stacks of merged confocal Z sections. Yellow arrowheads indicate the vertex; white arrows indicate axonal projection in the aboral ectoderm; numbers on the top right corner indicate the number of larvae analyzed and the number of larvae with similar staining patterns. M, Mouth

Although the data described in this last section result from experiments performed on larvae originating from two independent biological replicates and thus are preliminary results, several points can be made. First, there is a clear indication that the *Sp-An* neuropeptide is involved in neuronal growth and differentiation due to the nervous system impairment caused by its knockdown and possibly skeletal growth based on the larval vertex defects observed in *Sp-An* morphants. Interestingly, a role in promoting neuronal differentiation and growth would be in line with the role of neuropeptides in vertebrates, shown to act as neurotrophic factors necessary for proper neuronal development (Strand et al. 1991, Gozes and

Brenneman 1993, Gozes et al. 1997, Bakos et al. 2016). Finally, both *Sp-Pdx1* and *Sp-An* knockdowns lead to the downregulation of Th in the PPE neurons and a reduced number of apical plate serotonergic neurons. This observation implies that *Sp-Pdx1* induces neuronal differentiation through the neurotrophic effect of the *Sp-An* neuropeptide. Nonetheless, future studies addressing the role of An are needed before a definite conclusion can be drawn.

Chapter 6

General discussion

In this final chapter, I discuss the methodologies used, the results obtained, and the future perspectives of this work.

6.1. ScRNA-seq is able to reconstruct sea urchin cell types.

The main goal of this project is to unravel the cellular diversity during embryonic and larval development of the sea urchin *S. purpuratus*. To do so, I successfully developed and applied a single cell RNA sequencing pipeline that managed to reconstruct the cell types of one embryonic and two larval stages.

The success of this pipeline lies in many aspects that are linked to the use of sea urchin embryo/larva as a model system. First and foremost, sea urchins can produce billions of synchronized developing embryos, meaning an enormous amount of starting material. This starting material can be easily dissociated into single cells via a gentle enzyme-free dissociation protocol developed by me and colleagues that produces intact single cells in less than 20 min. Another advantage comes from using the *S. purpuratus* species, whose genome, transcriptome, and gene annotations are improved continuously. Last but not least, decades of research on *S. purpuratus* and sea urchin development, in general, led to the identification of several complex molecular specification and differentiation mechanisms at a gene regulatory level and the generation of one of the most detailed gene expression libraries.

Combining scRNA-seq with *in situ* hybridization, immunohistochemistry, and decades of knowledge regarding cell type molecular signature, I was able to successfully identify and map the computationally produced putative broad cell types on specific embryonic and larval territories. Furthermore, these putative broad cell types can be sub-divided into multiple subpopulations that can be traced to distinct embryonic and larval domains increasing the resolution in cellular composition complexity. Moreover, cross-stage integration of an embryonic and a larval dataset reveals high transcriptomic conservation of cell types despite morphological and physiological differences of the two developmental stages. Additional evidence of such conservation comes from the gene regulatory analysis performed in this thesis, in which, for instance, the pre-gastrula gene regulatory network components are found to be active in the pluteus skeletal cells.

ScRNA-seq is able to identify complex and dynamic gene expression patterns that were previously undetectable by *in situ* hybridization, as well as novel gene expression domains of well-described gene markers that suggest a new function.

Nevertheless, the most crucial part in unraveling the physiological mechanisms and evolution of a given cell type is deciphering its regulatory identity. The regulatory state analysis of the recognized putative cell types in respect to their developmental origins of the early pluteus stage shows that there are only a few transcription factors that are germ layer-specific, implying that either differentiation depends on a subset of rewired transcription factors or that common regulatory mechanisms promote differentiation in diverse germ layers. For instance, the gene regulatory module guiding the homing of the mesodermally derived small micromeres into the coelomic pouch is predicted by the scRNA-seq analysis to be also active in the ectodermally derived apical organ region.

Overall, the information provided by the single cell RNA sequencing data generated in this PhD thesis is a tremendous resource for identifying the cell types of a non-chordate deuterostome and if combined with other omics approaches, such as ATAC-seq, scATAC-seq, differential RNA-seq, Chip-seq, can result in drafting novel GRNs.

6.2. ScRNA-seq reveals unprecedented neuronal diversity

According to the analysis presented in this thesis, scRNA-seq can be used as a tool to deeply characterize the components of the sea urchin embryo and larva nervous system as well as to identify novel neuronal subtypes.

The results of the analysis performed in this PhD thesis highlight that different pre-neuronal and neuronal populations show a high degree of transcriptional similarity, despite the already, in some cases, demonstrated distinct developmental origins and functions, allowing their grouping in putative broad cell types. This is most probably a common feature across the various putative broad cell types of the developing sea urchin embryo and larva as indicated by the observation according to which multiple putative broad cell types (e.g., immune and skeletal) can be subclustered into subtypes that can be traced to different regions of the organism.

Such subclustering analyses revealed a higher number of neuronal diversity and complexity than previously described, which I was able to map to distinct regions of the *S. purpuratus* embryo and larva, always taking advantage of the thorough studies carried out in the sea urchin in the past. Furthermore, single cell RNA

sequencing can be used as a tool to address the evolution of the nervous system as well as to unravel the functional relationships of the sea urchin larva nervous system with its cell type targets. For instance, as shown in this thesis, scRNA-seq can provide useful information on the spatial expression of the major components involved in neurotransmission pathways and the machinery utilized by sea urchin neurons, increasing our understanding of the role and evolution of the nervous system. Examples of this are identifying possible cell type targets of the serotonergic and dopaminergic signaling pathways and the revelation of the interplay of non-neuronal types controlling the neurotransmission process. For instance, scRNA-seq verified the non-neuronal mediated GABA-ergic singling and revealed a possible role of the digestive tract in the inactivation of the acetylcholine mediated signaling by selectively expressing the enzyme that breaks it down. Moreover, scRNA-seq, in combination with improved gene annotation, reveals the presence of signaling systems so far never described and considered to be absent in sea urchins, such as the endocannabinoid signaling and its possible involvement both in neuronal and non-neuronal communication, as well as the signaling mediated by endogenously produced melatonin.

Although single cell RNA sequencing is an enormous resource for getting a glimpse of the regulatory landscape operating within cell types, its real power is revealed when combined with other approaches such as traditional molecular biology visualization and gene expression perturbation techniques. As a proof of principle case study, I used the scRNA-seq derived information on the expression domains of two nitric oxide synthase genes producing the unconventional neurotransmitter NO to speculate what its role might be. Then, after blocking its production and assessing the phenotype, I used the scRNA-seq data to understand the mechanism by which NO exhibits its function. This led to identifying its role in larval morphogenesis (perturbation analysis) and predicting in which cell types it establishes its function through canonical signaling (single cell transcriptomics).

ScRNA-seq is also able to highlight transcriptional similarities that can be used in unraveling the evolution of a cell type. In the case of the sea urchin anterior neuroectoderm/apical plate pre-neuronal domains, scRNA-seq verifies the presence of a gene regulatory compartmentalization as previously proposed, which is in line with the high neuronal diversity that seems to be associated with them as revealed by this work. Furthermore, the same analysis shows that the anterior/apical

domain of an invertebrate uses a similar gene toolkit to the one guiding the specification of the vertebrate anterior neural plate that will give rise to the forebrain, supporting the hypothesis that these two regions have a common evolutionary history.

Overall, scRNA-seq is a powerful tool to discover novel neuronal cell types, and the tremendous amount of information that derives from the scRNA-seq data contributes to the understanding of their function and cross-talk with non-neuronal cell types.

6.3. Distinct sea urchin cell types employ a pancreatic-like gene toolkit

The emergence of the pancreas as a distinct organ is a feature solely associated with the vertebrate lineage. However, several specialized cell types with pancreatic-like characteristics have been found in various invertebrate taxa, suggesting that the functional cell units of vertebrate pancreas predate its evolutionary emergence of a distinct organ. Such cell types have been found in the *S. purpuratus* pluteus larva in distinct digestive tract domains. Moreover, based on the many morphological, physiological and molecular similarities between neurons and β-cells, it has been hypothesized that one of the functional units of the endocrine pancreas (β-cells) arose during evolution by co-opting a preexisting neuronal gene regulatory program. Based on the above, one of the ultimate goals of this project is to identify at single cell resolution which are the cell types of the sea urchin larva that have a pancreatic-like molecular signature and whether such a signature is found in sea urchin neurons.

The scRNA-seq analysis presented in this work reveals that key genes involved in the organogenesis and function of the vertebrate pancreas are enriched in distinct larval cell types. In particular, an exocrine pancreas-like molecular signature is found restricted to a single putative broad cell type. Simultaneously, genes involved in endocrine pancreas formation are expressed in two endodermal cell types, known to give rise to insulinergic cells later in development and the nervous system. Overall, it is evident that the digestive tract of the 3 dpf *S. purpuratus* larva contains cells utilizing distinct exocrine and endocrine genetic programs. The exocrine pancreas-like cells appear fully differentiated as judged by the expression of terminal differentiation markers (amylase, lipase, proteinase, carboxypeptidase). In

contrast, the endocrine pancreas-like cell types are possibly still differentiating, in line with previous reports and based on the absence of insulin transcripts from the scRNA-seq datasets (Perillo 2013).

The endocrine pancreas-like molecular signature found in the nervous system of the sea urchin larva contains several shared elements with the endocrine-like cells identified in the gut, including the key transcription factors *FoxA, Pdx1, NeuroD1, Mnx,* and *Rfx3*, being evidence of a possible gene regulatory wiring co-option. Further analysis shows that the endocrine-pancreas molecular fingerprint is widespread throughout the different neuronal populations, although enriched in distinct neuronal types. Furthermore, molecular perturbation analysis of two neuronal types expressing the sea urchin homolog of the mammalian transcription factor *Pdx1*, a master gene controlling pancreas and β-endocrine cell formation, shows that *Sp-Pdx1* is essential for proper differentiation of the neuronal types in which it is expressed. Further analysis of an *Sp-Pdx1* positive neuronal population, in which also the sea urchin homolog (*Sp-Brn1/2/4*) of the mammalian α-endocrine cell-specific marker *Brn4* is co-expressed, resulted in the identification of 20 genes whose mammalian homologs are involved in the differentiation and function of pancreatic β-cells. Moreover, combining the information coming from both scRNA-seq and differential RNA-seq performed on *Sp-Pdx1* morphants, I was able to reconstruct a putative GRN wiring the *Sp-Pdx1/Sp-Brn1/2/4* neurons. The resulting GRN depicts the crucial role of *Sp-Pdx1* in acquiring a neurosecretory fate, while preliminary results suggest that the production of the *Sp-Pdx1* dependent neuropeptide *Sp-An* is also involved in this process, possibly acting as a neurotrophic factor.

Overall, the data produced suggest that specialized neurons with an endocrine pancreas-like molecular signature are present in the sea urchin larva and employ an *Sp-Pdx1* orchestrated genetic program similar to the one used by mammalian endocrine cells. Although such a role for *Pdx1* has not been revealed by scRNA-seq studies in other organisms yet, it has been recently shown that the adult oyster insulin gene is a likely direct target of *Pdx1*, in a tissue that transcriptomically seems to be pancreas related (Xu et al. 2021).

The sea urchin *Sp-Pdx1/Sp-Brn1/2/*4 neurons express a high number of key genes that are necessary for the endocrine pancreas differentiation and function in the vertebrates. These shared molecular features of Sp-Pdx1/Sp-Brn1/2/4 neurons and

pancreatic cells represent features of a cell type present in the deuterostome ancestor. What was the role of these cells and what happened in the evolution of the descendant lineages?

One scenario is that the ancestral deuterostomes already possessed a *Sp-Pdx1/Sp-Brn1/2/4* neuronal cell type similar to the one found in today's sea urchin, and that the regulatory program of this cell type was then co-opted into the endodermal lineage in the vertebrates. The co-option and incorporation of gene regulatory elements to different developmental or morphogenetic context has been postulated before and has the potential to give rise to diverse cell types (Martik and McClay, 2015, Morgulis et al., 2019, Cary et al., 2020). In particular, it has been hypothesized that β pancreatic cells arose during evolution by co-option of a preexisting neuronal cell type program into the pancreas developmental lineage (Arntfield and van der Kooy, 2011, Perillo et al., 2018) based on the many physiological, morphological and molecular features endocrine pancreatic cells share with neurons (Alpert et al., 1988, Eberhard, 2013).

A second possibility is that both ectodermal neurons and endodermal pancreatic cells are direct evolutionary descendants of the same precursor cell (Arendt, 2008). This explanation would avoid mean the existence of a direct evolutionary link between cell types derived from different germ layers - ectoderm versus endoderm. Moreover, a similar link between digestive and neuronal-type expression profiles has been reported repeatedly. In the sea anemone *Nematostella vectensis*, pharyngeal cells give rise to both digestive cells and neurons (Steinmetz et al., 2017), and neuronal cell types and secretory gland cells exhibit related molecular profiles (Sebe-Pedros et al., 2018). In fact, based on these findings a close evolutionary link between endoderm and ectoderm has been postulated (Steinmetz et al., 2017). Finally, motor neurons in the vertebrate ventral neural tube and pancreatic islet cells share a specific combination of transcription factors (reviewed in Arendt, 2021). Overall both hypotheses are in line with the fact that at least subsets of neuron types in animal nervous systems may have originated from cell population with both digestive and communicative functions (Arendt et al., 2015; Arendt, 2021).

6.4. Future directions

6.4.1. Single cell transcriptomics

The single cell libraries constructed for the 2 dpf gastrula and 3 dpf early pluteus stages are able to fully reconstruct the cell types present at these two developmental stages. This is mainly because of the large number of different biological and technical replicates produced. In the case of the 5 dpf pluteus dataset, which is a result of the integration of libraries coming from 2 biological replicates, although I was able to identify most of the cell types expected to be present, I noticed less complexity regarding the endodermally derived cell types when compared to the 3 dpf dataset. This observation, together with the higher number of cells that constitute the 5 dpf larva, suggest that more replicates are needed to fully reconstruct all the cell types present at this stage. Moreover, the current data concern only the aforementioned developmental time-points. To reconstruct developmental trajectories, depicting dynamic gene regulatory mechanisms active during cell type specification and differentiation as well as to trace the developmental origins of the identified in this thesis neuronal population, more single cell RNA sequencing for closely related developmental stages is needed.

6.4.2. *In vivo* validation of scRNA-seq predictions

The majority of the topics addressed throughout this PhD thesis are based on gene expression pattern predictions deriving from the scRNA-seq analysis and is paired with *in vivo* validation through *in situ* hybridization and/or immunohistochemistry and/or functional experiments. Nonetheless, there are cases such as the previously not described in sea urchin endocannabinoid, melatonin, and nitric oxide signalings, for which such *in vivo* validations are still needed before reaching a safe conclusion.

6.4.3. Microinjection of morpholino antisense oligonucleotides

Perturbation of gene expression is a key step to identify the role of a given gene. The role of selected genes was assessed in this work by microinjection of specific morpholino antisense oligonucleotides into fertilized sea urchin eggs. The best practice for reaching a safe conclusion on the effect of a morpholino and thus the

function of a given gene is performing the experiments on embryos/larvae coming from at least three different biological replicates. Therefore, the experiments concerning the morpholino injections causing the knockdown of Sp-NosA, Sp-NosB, Sp-NosA/B (carried out on larvae coming from one biological replicate) Sp-An, (performed on larvae originating from two biological replicates) need to be repeated.

6.4.4. Gene regulatory network reconstruction

The molecular fingerprint revealed by the single cell transcriptomics is an overview of the regulatory interactions active in a given cell type. To decipher the hierarchy of these gene regulatory interactions leading to the acquisition and maintenance of a cell fate, perturbation of gene expression and *in vivo* validations is crucial. Such further analysis is required to fully reconstruct the anterior neuroectoderm/apical plate and the PPE neurons gene regulatory networks. Further information on whether observed gene interactions are direct or indirect will be provided by incorporating information coming from bulk ATAC-seq data generated by the lab in which I carried out this project.

6.4.5. Ultrastructural imaging of the whole sea urchin larva and registration of gene expression atlas

Recent technological advancements allow the simultaneous use of electron microscopy (EM) and 3D reconstruction to investigate biological structures (Titze and Genoud 2016). A combination of this technology with the information obtained from the single cell transcriptomics would allow the reconstruction of a comprehensive atlas depicting the *S. purpuratus* cell types both at a morphological/ultrastructural and molecular level (gene expression).

6.5. Conclusions

The aims and goals set for this PhD work, indicated in the introduction chapter, have been successfully achieved, and the main achievements are listed as follows:

- Successful development and application of protocols for sea urchin whole embryo/larva dissociation and single cell RNA sequencing.

- Identification and reconstruction of the *S. purpuratus* embryonic and larval cell types at a single cell resolution.
- Identification of new gene markers and prediction of novel gene regulatory module functions.
- Discovery of unprecedented neuronal diversity and thorough characterization of neuronal signalings.
- Characterization of two *Nitric oxide synthase* genes and unraveling a possible role of NO mediated signaling in sea urchin.
- Identification of ANE/apical plate subtypes and of a conserved gene toolkit.
- Identification of cell types with pancreatic-like molecular signature at a single cell resolution.
- Reporting a role of *Sp-Pdx1* in neuronal differentiation and acquisition of neurosecretory fate.

Overall, the data produced for this PhD thesis contribute to expanding our understanding of cell type and nervous system evolution and set a solid foundation for future research.

Publications

1. Margherita Perillo, Periklis Paganos, Teresa Mattiello, Maria Cocurullo, Paola Oliveri, Maria Ina Arnone, 2018, New neuronal subtypes with a "pre-pancreatic" signature in the sea urchin *Strongylocentrotus purpuratus*, Frontiers in endocrinology, https://doi.org/10.3389/fendo.2018.00650

   Contribution: Acquisition, analysis, and interpretation of data.

2. Enrico D'Aniello, Periklis Paganos, Evgeniya Anishchenko, Salvatore D'Aniello, Maria Ina Arnone, 2020, Comparative neurobiology of biogenic amines in animal models in deuterostomes, Frontiers in Ecology and Evolution, https://doi.org/10.3389/fevo.2020.587036

   Contribution: Writing and revising of the manuscript and figures.

3. Flora Rendell-Bhatti, Periklis Paganos, Anna Pouch, Christopher Mitchell, Salvatore D'Aniello, Brendan J Godley, Ksenia Pazdro, Maria Ina Arnone, Eva Jimenez-Guri, 2020, Developmental toxicity of plastic leachates on the sea urchin *Paracentrotus lividus*, Environmental Pollution, https://doi.org/10.1016/j.envpol.2020.115744

   Contribution: Investigation, formal analysis, visualization, writing, review & editing

4. Margherita Perillo, Periklis Paganos, Maxwell Spurrell, Maria Ina Arnone, Gary M Wessel, 2021, Methodology for Whole Mount and Fluorescent RNA *In Situ* Hybridization in Echinoderms: Single, Double, and Beyond, Developmental Biology of the Sea Urchin and Other Marine Invertebrates, https://doi.org/0.1007/978-1-0716-0974-3_12

   Contribution: Protocol adaptation, writing of the chapter.

5. Ioannis Tsironis, Periklis Paganos, Georgia Gouvi, Panagiotis Tsimpos, Andriana Stamopoulou, Maria Ina Arnone, Constantin N Flytzanis, 2021, Coup-TF: A maternal factor essential for differentiation along the embryonic axes in the sea urchin *Paracentrotus lividus*, Developmental Biology, https://doi.org/10.1016/j.ydbio.2020.12.012

Contribution: First co-author, designing, writing, acquisition, analysis and interpretation of data

6. Jeffrey R Thompson, Periklis Paganos, Giovanna Benvenuto, Maria Ina Arnone, Paola Oliveri, under revision, Post-metamorphic skeletal growth in the sea urchin *Paracentrotus lividus* and implications for body plan evolution, Evo-Devo, https://doi.org/10.1186/s13227-021-00174-1

   Contribution: Acquisition of data, imaging, review & editing

7. Periklis Paganos, Danila Voronov, Jacob Musser, Detlev Arendt, Maria Ina Arnone, Single cell RNA sequencing of the *Strongylocentrotus purpuratus* larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. bioRxiv 2021.03.16.435574; doi: https://doi.org/10.1101/2021.03.16.435574

   Contribution: Conceptualization, writing, acquisition, analysis and interpretation of data

Non-Book component

Non-book component contains detailed results described in Chapters 3, 4, and 5. The non-book component has:

- Scripts used to analyze the single cell data and perform the clustering analysis (Sp_2dpf_clustering_analysis, Sp_3dpf_clustering_analysis, Sp_5dpf_clustering_analysis, Sp_2_3_dpf_clustering_analysis).
- Script used to perform the subclustering analysis (Subclustering analysis).
- Tables containing the marker genes and the total genes expressed per cell types across the three developmental time-points analyzed in this thesis (Sp_2dpf_all_genes, Sp_2dpf_marker_genes, Sp_3dpf_all_genes, Sp_3dpf_marker_genes, Sp_5dpf_all_genes, Sp_5dpf_marker_genes).
- A table containing the total genes expressed per neuronal population at 3 dpf pluteus stage (Sp_3 dpf neuronal subclustering_all genes).
- Tables that contain the differentially expressed genes as a result of Sp-Pdx1 knockdown at 3 dpf and the specific targets of Sp-Pdx1 in the PPE neurons (Sp-Pdx1 differential RNA-seq, Differentially expressed genes in the PPE neurons).
- A table containing additional gene annotations of all the genes expressed in all datasets (Sp_genes_annotation_NCBI_GO_PFAM_FUNCTIONAL).

Bibliography

Adams, D. K., M. A. Sewell, R. C. Angerer and L. M. Angerer (2011). "Rapid adaptation to food availability by a dopamine-mediated morphogenetic response." Nat Commun **2**: 592.

Adams, N. L., A. Heyland , L. L. Rice, K. R. Foltz (2019)." Chapter 1 - Procuring animals and culturing of eggs and embryos." Methods in Cell Biology **150:** 3-46.

Afelik, S., Y. Chen and T. Pieler (2006). "Combined ectopic expression of Pdx1 and Ptf1a/p48 results in the stable conversion of posterior endoderm into endocrine and exocrine pancreatic tissue." Genes Dev **20**(11): 1441-1446.

Ahlgren, U., J. Jonsson and H. Edlund (1996). "The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice." Development **122**(5): 1409-1416.

Aiello, V., A. Moreno-Asso, J. M. Servitja and M. Martin (2014). "Thyroid hormones promote endocrine differentiation at expenses of exocrine tissue." Exp Cell Res **322**(2): 236-248.

Ait-Lounis, A., C. Bonal, Q. Seguin-Estevez, C. D. Schmid, P. Bucher, P. L. Herrera, B. Durand, P. Meda and W. Reith (2010). "The transcription factor Rfx3 regulates beta-cell differentiation, function, and glucokinase expression." Diabetes **59**(7): 1674-1685.

Alpert, S., D. Hanahan and G. Teitelman (1988). "Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply a relationship with neurons." Cell **53**(2): 295-308.

Amisten, S., M. Neville, R. Hawkes, S. J. Persaud, F. Karpe and A. Salehi (2015). "An atlas of G-protein coupled receptor expression and function in human subcutaneous adipose tissue." Pharmacol Ther **146**: 61-93.

Andreakis, N., S. D'Aniello, R. Albalat, F. P. Patti, J. Garcia-Fernandez, G. Procaccini, P. Sordino and A. Palumbo (2011). "Evolution of the nitric oxide synthase family in metazoans." Mol Biol Evol **28**(1): 163-179.

Andrikou, C., E. Iovene, F. Rizzo, P. Oliveri and M. I. Arnone (2013). "Myogenesis in the sea urchin embryo: the molecular fingerprint of the myoblast precursors." Evodevo **4**(1): 33.

Andrikou, C., C. Y. Pai, Y. H. Su and M. I. Arnone (2015). "Logics and properties of a genetic regulatory program that drives embryonic muscle development in an echinoderm." Elife **4**.

Angerer, L. M., S. Yaguchi, R. C. Angerer and R. D. Burke (2011). "The evolution of nervous system patterning: insights from sea urchin development." Development **138**(17): 3613-3623.

Anishchenko, E., M. I. Arnone and S. D'Aniello (2018). "SoxB2 in sea urchin development: implications in neurogenesis, ciliogenesis and skeletal patterning." Evodevo **9**: 5.

Annona, G., F. Caccavale, J. Pascual-Anaya, S. Kuratani, P. De Luca, A. Palumbo and S. D'Aniello (2017). "Nitric Oxide regulates mouth development in amphioxus." Sci Rep **7**(1): 8432.

Annunziata, R. and M. I. Arnone (2014). "A dynamic regulatory network explains ParaHox gene control of gut patterning in the sea urchin." Development **141**(12): 2462-2472.

Annunziata, R., M. Perillo, C. Andrikou, A. G. Cole, P. Martinez and M. I. Arnone (2014). "Pattern and process during sea urchin gut morphogenesis: the regulatory landscape." Genesis **52**(3): 251-268.

Araujo, D. J., K. Tjoa and K. Saijo (2019). "The Endocannabinoid System as a Window Into Microglial Biology and Its Relationship to Autism." Front Cell Neurosci **13**: 424.

Arda, H. E., C. M. Benitez and S. K. Kim (2013). "Gene regulatory networks governing pancreas development." Dev Cell **25**(1): 5-13.

Arendt, D. (2008). "The evolution of cell types in animals: emerging principles from molecular studies". Nat Rev Genet, **9**, 868–82

Arendt, D. (2021). "Elementary nervous systems". Phil. Trans. R. Soc. B 376: 20200347

Arendt, D., J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler and G. P. Wagner (2016). "The origin and evolution of cell types." Nat Rev Genet **17**(12): 744-757.

Arkhipova, V., B. Wendik, N. Devos, O. Ek, B. Peers and D. Meyer (2012). "Characterization and regulation of the hb9/mnx1 beta-cell progenitor specific enhancer in zebrafish." Dev Biol **365**(1): 290-302.

Arnone, M. I. and E. H. Davidson (1997). "The hardwiring of development: organization and function of genomic regulatory systems." Development **124**(10): 1851-1864.

Arnone, M. I., I. J. Dmochowski and C. Gache (2004). "Using reporter genes to study cis-regulatory elements." Methods Cell Biol **74**: 621-652.

Arntfield, M. E. and D. van der Kooy (2011). "beta-Cell evolution: How the pancreas borrowed from the brain: The shared toolbox of genes expressed by neural and pancreatic endocrine cells may reflect their evolutionary relationship." Bioessays **33**(8): 582-587.

Baek, S. and I. Lee (2020). "Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation." Computational and Structural Biotechnology Journal **18**: 1429-1439.

Bakos, J., M. Zatkova, Z. Bacova and D. Ostatnikova (2016). "The Role of Hypothalamic Neuropeptides in Neurogenesis and Neuritogenesis." Neural Plast **2016**: 3276383.

Baran-Gale J., T. Chandra, K. Kirschner. (2018). "Experimental design for single-cell RNA sequencing", Briefings in Functional Genomics, 17(4), 233–239.

Baron, M., A. Veres, S. L. Wolock, A. L. Faust, R. Gaujoux, A. Vetere, J. H. Ryu, B. K. Wagner, S. S. Shen-Orr, A. M. Klein, D. A. Melton and I. Yanai (2016). "A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure." Cell Syst **3**(4): 346-360 e344.

Barsi, J. C., E. Li and E. H. Davidson (2015). "Geometric control of ciliated band regulatory states in the sea urchin embryo." Development **142**(5): 953-961.

Batterham, R. L., C. W. Le Roux, M. A. Cohen, A. J. Park, S. M. Ellis, M. Patterson, G. S. Frost, M. A. Ghatei and S. R. Bloom (2003). "Pancreatic polypeptide reduces appetite and food intake in humans." J Clin Endocrinol Metab **88**(8): 3989-3992.

Baudonnat, M., A. Huber, V. David and M. E. Walton (2013). "Heads for learning, tails for memory: reward, reinforcement and a role of dopamine in determining behavioral relevance across multiple timescales." Front Neurosci **7**: 175.

Beccari, L., R. Marco-Ferreres and P. Bovolenta (2013). "The logic of gene regulatory networks in early vertebrate forebrain patterning." Mech Dev **130**(2-3): 95-111.

Becher, A., A. Drenckhahn, I. Pahner, M. Margittai, R. Jahn and G. Ahnert-Hilger (1999). "The synaptophysin-synaptobrevin complex: a hallmark of synaptic vesicle maturation." J Neurosci **19**(6): 1922-1931.

Belgacem, Y. H., A. M. Hamilton, S. Shim, K. A. Spencer and L. N. Borodinsky (2016). "The Many Hats of Sonic Hedgehog Signaling in Nervous System Development and Disease." J Dev Biol **4**(4).
Belle, J. I. and D. G. DeNardo (2019). "A Single-Cell Window into Pancreas Cancer Fibroblast Heterogeneity." Cancer Discov **9**(8): 1001-1002.

Ben-Tabou de-Leon, S., Y. H. Su, K. T. Lin, E. Li and E. H. Davidson (2013). "Gene regulatory control in the sea urchin aboral ectoderm: spatial initiation, signaling inputs, and cell fate lockdown." Dev Biol **374**(1): 245-254.

Benadiba, C., D. Magnani, M. Niquille, L. Morle, D. Valloton, H. Nawabi, A. Ait-Lounis, B. Otsmane, W. Reith, T. Theil, J. P. Hornung, C. Lebrand and B. Durand (2012). "The ciliogenic transcription factor RFX3 regulates early midline distribution of guidepost neurons required for corpus callosum development." PLoS Genet **8**(3): e1002606.

Berger, C. and D. Zdzieblo (2020). "Glucose transporters in pancreatic islets." Pflugers Arch **472**(9): 1249-1272.

Bisgrove, B. W. and R. D. Burke (1987). "Development of the nervous system of the pluteus larva of Strongylocentrotus droebachiensis." Cell and Tissue Research **248**(2): 335-343.

Blader, P., C. S. Lam, S. Rastegar, R. Scardigli, J. C. Nicod, N. Simplicio, C. Plessy, N. Fischer, C. Schuurmans, F. Guillemot and U. Strahle (2004). "Conserved and acquired features of neurogenin1 regulation." Development **131**(22): 5627-5637.

Boucrot, E., S. Saffarian, R. Zhang and T. Kirchhausen (2010). "Roles of AP-2 in clathrin-mediated endocytosis." PLoS One **5**(5): e10597.

Boutin, C., O. Hardt, A. de Chevigny, N. Core, S. Goebbels, R. Seidenfaden, A. Bosio and H. Cremer (2010). "NeuroD1 induces terminal neuronal differentiation in olfactory neurogenesis." Proc Natl Acad Sci U S A **107**(3): 1201-1206.

Briggs, E. and G. M. Wessel (2006). "In the beginning… Animal fertilization and sea urchin development." Developmental Biology **300**(1): 15-26.

Broillet, M. C. (1999). "S-nitrosylation of proteins." Cell Mol Life Sci **55**(8-9): 1036-1042.

Buckle, A., R. S. Nozawa, D. A. Kleinjan and N. Gilbert (2018). "Functional characteristics of novel pancreatic Pax6 regulatory elements." Hum Mol Genet **27**(19): 3434-3448.

Burke, R. D. (1978). "The structure of the nervous system of the pluteus larva of Strongylocentrotus purpuratus." Cell Tissue Res **191**(2): 233-247.

Burke, R. D., L. M. Angerer, M. R. Elphick, G. W. Humphrey, S. Yaguchi, T. Kiyama, S. Liang, X. Mu, C. Agca, W. H. Klein, B. P. Brandhorst, M. Rowe, K. Wilson, A. M. Churcher, J. S. Taylor, N. Chen, G. Murray, D. Wang, D. Mellott, R. Olinski, F. Hallbook and M. C. Thorndyke (2006). "A genomic view of the sea urchin nervous system." Dev Biol **300**(1): 434-460.

Burke, R. D., D. J. Moller, O. A. Krupke and V. J. Taylor (2014). "Sea urchin neural development and the metazoan paradigm of neurogenesis." Genesis **52**(3): 208-221.

Burke, R. D., L. Osborne, D. Wang, N. Murabe, S. Yaguchi and Y. Nakajima (2006). "Neuron-specific expression of a synaptotagmin gene in the sea urchin Strongylocentrotus purpuratus." J Comp Neurol **496**(2): 244-251.

Burkhardt, P. and S. G. Sprecher (2017). "Evolutionary origin of synapses and neurons - Bridging the gap." Bioessays **39**(10).

Burlison, J. S., Q. Long, Y. Fujitani, C. V. Wright and M. A. Magnuson (2008). "Pdx-1 and Ptf1a concurrently determine fate specification of pancreatic multipotent progenitor cells." Dev Biol **316**(1): 74-86.

Butler, A., P. Hoffman, P. Smibert, E. Papalexi and R. Satija (2018). "Integrating single-cell transcriptomic data across different conditions, technologies, and species." Nat Biotechnol **36**(5): 411-420.

Cao, C., L. A. Lemaire, W. Wang, P. H. Yoon, Y. A. Choi, L. R. Parsons, J. C. Matese, W. Wang, M. Levine and K. Chen (2019). "Comprehensive single-cell transcriptome lineages of a proto-vertebrate." Nature **571**(7765): 349-354.

Carter, H. F., J. R. Thompson, M. R. Elphick and P. Oliveri (2021). "The development and neuronal complexity of bipinnaria larvae of the sea star Asterias rubens." bioRxiv: 2021.2001.2004.425292.

Castellano, I., E. Ercolesi and A. Palumbo (2014). "Nitric oxide affects ERK signaling through down-regulation of MAP kinase phosphatase levels during larval development of the ascidian Ciona intestinalis." PLoS One **9**(7): e102907.

Castillo, P. E., T. J. Younts, A. E. Chavez and Y. Hashimotodani (2012). "Endocannabinoid signaling and synaptic function." Neuron **76**(1): 70-81.

Cheers, M. S. and C. A. Ettensohn (2004). "Rapid microinjection of fertilized eggs." Methods Cell Biol **74**: 287-310.

Chen, K. and N. Rajewsky (2007). "The evolution of gene regulation by transcription factors and microRNAs." Nat Rev Genet **8**(2): 93-103.

Chen, K., A. Richlitzki, D. E. Featherstone, M. Schwarzel and J. E. Richmond (2011). "Tomosyn-dependent regulation of synaptic transmission is required for a late phase of associative odor memory." Proc Natl Acad Sci U S A **108**(45): 18482-18487.

Chestnut, B., S. Casie Chetty, A. L. Koenig and S. Sumanas (2020). "Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle." Nat Commun **11**(1): 2796.

Cho, B., S. H. Yoon, D. Lee, F. Koranteng, S. G. Tattikota, N. Cha, M. Shin, H. Do, Y. Hu, S. Y. Oh, D. Lee, A. Vipin Menon, S. J. Moon, N. Perrimon, J. W. Nam and J. Shim (2020). "Single-cell transcriptome maps of myeloid blood cell lineages in Drosophila." Nat Commun **11**(1): 4483.

Churchill, A. J., G. D. Gutierrez, R. A. Singer, D. S. Lorberbaum, K. A. Fischer and L. Sussel (2017). "Genetic evidence that Nkx2.2 acts primarily downstream of Neurog3 in pancreatic endocrine lineage development." Elife **6**.

Coffman, J. A., J. J. McCarthy, C. Dickey-Sims and A. J. Robertson (2004). "Oral-aboral axis specification in the sea urchin embryo II. Mitochondrial distribution and redox state contribute to establishing polarity in Strongylocentrotus purpuratus." Dev Biol **273**(1): 160-171.

Cole, A. G. and M. I. Arnone (2009). "Fluorescent in situ hybridization reveals multiple expression domains for SpBrn1/2/4 and identifies a unique ectodermal cell type that co-expresses the ParaHox gene SpLox." Gene Expr Patterns **9**(5): 324-328.

Cole, A. G., F. Rizzo, P. Martinez, M. Fernandez-Serra and M. I. Arnone (2009). "Two ParaHox genes, SpLox and SpCdx, interact to partition the posterior endoderm in the formation of a functional gut." Development **136**(4): 541-549.

Colizzi, E. S., R. M. Vroomans and R. M. Merks (2020). "Evolution of multicellularity by collective integration of spatial information." Elife **9**.

Cooper, G. A. and S. A. West (2018). "Division of labour and the evolution of extreme specialization." Nat Ecol Evol **2**(7): 1161-1167.

Cottone, E., V. Pomatto and P. Bovolin (2013). "Role of the endocannabinoid system in the central regulation of nonmammalian vertebrate reproduction." Int J Endocrinol **2013**: 941237.

Cristino, L., V. Guglielmotti, A. Cotugno, C. Musio and S. Santillo (2008). "Nitric oxide signaling pathways at neural level in invertebrates: functional implications in cnidarians." Brain Res **1225**: 17-25.

D'Amico, L. A., D. Boujard and P. Coumailleau (2013). "The neurogenic factor NeuroD1 is expressed in post-mitotic cells during juvenile and adult Xenopus neurogenesis and not in progenitor or radial glial cells." PLoS One **8**(6): e66487.

D'Aniello, E., P. Paganos, E. Anishchenko, S. D'Aniello and M. I. Arnone (2020). "Comparative Neurobiology of Biogenic Amines in Animal Models in Deuterostomes." Frontiers in Ecology and Evolution **8**(322).

Dalgin, G. and V. E. Prince (2015). "Differential levels of Neurod establish zebrafish endocrine pancreas cell fates." Dev Biol **402**(1): 81-97.

Davidson, E. H. (2006). "The regulatory genome : gene regulatory networks in development and evolution." from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=166274.

Davidson, E. H. and D. H. Erwin (2006). "Gene regulatory networks and the evolution of animal body plans." Science **311**(5762): 796-800.

Davidson, P. L., H. Guo, L. Wang, A. Berrio, H. Zhang, Y. Chang, A. L. Soborowski, D. R. McClay, G. Fan and G. A. Wray (2020). "Chromosomal-Level Genome Assembly of the Sea Urchin Lytechinus variegatus Substantially Improves Functional Genomic Analyses." Genome Biol Evol **12**(7): 1080-1086.

Davie, K., J. Janssens, D. Koldere, M. De Waegeneer, U. Pech, L. Kreft, S. Aibar, S. Makhzami, V. Christiaens, C. Bravo Gonzalez-Blas, S. Poovathingal, G. Hulselmans, K. I. Spanier, T. Moerman, B. Vanspauwen, S. Geurs, T. Voet, J. Lammertyn, B. Thienpont, S. Liu, N. Konstantinides, M. Fiers, P. Verstreken and S. Aerts (2018). "A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain." Cell **174**(4): 982-998 e920.

de la Pompa, J. L., A. Wakeham, K. M. Correia, E. Samper, S. Brown, R. J. Aguilera, T. Nakano, T. Honjo, T. W. Mak, J. Rossant and R. A. Conlon (1997). "Conservation

of the Notch signalling pathway in mammalian neurogenesis." Development **124**(6): 1139-1148.

Decker, K., D. C. Goldman, C. L. Grasch and L. Sussel (2006). "Gata6 is an important regulator of mouse pancreas development." Dev Biol **298**(2): 415-429.

Diez Del Corral, R. and A. V. Morales (2017). "The Multiple Roles of FGF Signaling in the Developing Spinal Cord." Front Cell Dev Biol **5**: 58.

Ding, K., L. Zhang, T. Zhang, H. Yang and R. Brinkman (2019). "The Effect of Melatonin on Locomotor Behavior and Muscle Physiology in the Sea Cucumber Apostichopus japonicus." Front Physiol **10**: 221.

Doronzo, G., M. Viretto, I. Russo, L. Mattiello, L. Di Martino, F. Cavalot, G. Anfossi and M. Trovati (2011). "Nitric oxide activates PI3-K and MAPK signalling pathways in human and rat vascular smooth muscle cells: influence of insulin resistance and oxidative stress." Atherosclerosis **216**(1): 44-53.

Doyle, M. J. and L. Sussel (2007). "Nkx2.2 regulates beta-cell function in the mature islet." Diabetes **56**(8): 1999-2007.

Duboc, V., F. Lapraz, L. Besnardeau and T. Lepage (2008). "Lefty acts as an essential modulator of Nodal activity during sea urchin oral-aboral axis formation." Dev Biol **320**(1): 49-59.

Duboc, V., F. Lapraz, A. Saudemont, N. Bessodes, F. Mekpoh, E. Haillot, M. Quirin and T. Lepage (2010). "Nodal and BMP2/4 pattern the mesoderm and endoderm during development of the sea urchin embryo." Development **137**(2): 223-235.

Duboc, V., E. Rottinger, L. Besnardeau and T. Lepage (2004). "Nodal and BMP2/4 signaling organizes the oral-aboral axis of the sea urchin embryo." Dev Cell **6**(3): 397-410.

Eberhard, D. (2013). "Neuron and beta-cell evolution: learning about neurons is learning about beta-cells." Bioessays **35**(7): 584.

Ediger, B. N., A. Du, J. Liu, C. S. Hunter, E. R. Walp, J. Schug, K. H. Kaestner, R. Stein, D. A. Stoffers and C. L. May (2014). "Islet-1 Is essential for pancreatic beta-cell function." Diabetes **63**(12): 4206-4217.

Edlund, H. (2002). "Pancreatic organogenesis--developmental mechanisms and implications for therapy." Nat Rev Genet **3**(7): 524-532.

Efrat, S. and H. A. Russ (2012). "Making beta cells from adult tissues." Trends Endocrinol Metab **23**(6): 278-285.

Elphick, M. R. (2012). "The evolution and comparative neurobiology of endocannabinoid signalling." Philos Trans R Soc Lond B Biol Sci **367**(1607): 3201-3215.

Elphick, M. R. and R. Melarange (1998). "Nitric Oxide Function in an Echinoderm." Biol Bull **194**(3): 260-266.

Elphick, M. R., O. Mirabeau and D. Larhammar (2018). "Evolution of neuropeptide signalling systems." Journal of Experimental Biology **221**(3).

Elphick, M. R., Y. Satou and N. Satoh (2003). "The invertebrate ancestry of endocannabinoid signalling: an orthologue of vertebrate cannabinoid receptors in the urochordate Ciona intestinalis." Gene **302**(1-2): 95-101.

Enge, M., H. E. Arda, M. Mignardi, J. Beausang, R. Bottino, S. K. Kim and S. R. Quake (2017). "Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns." Cell **171**(2): 321-330 e314.

Ernst, S. G. (2011). "Offerings from an urchin." Dev Biol **358**(2): 285-294.

Esaulova, E., C. Cantoni, I. Shchukina, K. Zaitsev, R. C. Bucelli, G. F. Wu, M. N. Artyomov, A. H. Cross and B. T. Edelson (2020). "Single-cell RNA-seq analysis of human CSF microglia and myeloid cells in neuroinflammation." Neurol Neuroimmunol Neuroinflamm **7**(4).

Ettensohn, C. A. (2020). "The gene regulatory control of sea urchin gastrulation." Mech Dev: 103599.

Falkmer, S., E. Dafgard, M. el-Salhy, W. Engstrom, L. Grimelius and A. Zetterberg (1985). "Phylogenetical aspects on islet hormone families: a minireview with particular reference to insulin as a growth factor and to the phylogeny of PYY and NPY immunoreactive cells and nerves in the endocrine and exocrine pancreas." Peptides **6 Suppl 3**: 315-320.

Fernandez-Chacon, R., A. Konigstorfer, S. H. Gerber, J. Garcia, M. F. Matos, C. F. Stevens, N. Brose, J. Rizo, C. Rosenmund and T. C. Sudhof (2001). "Synaptotagmin I functions as a calcium regulator of release probability." Nature **410**(6824): 41-49.

Ferri, A. L., W. Lin, Y. E. Mavromatakis, J. C. Wang, H. Sasaki, J. A. Whitsett and S. L. Ang (2007). "Foxa1 and Foxa2 regulate multiple phases of midbrain dopaminergic neuron development in a dosage-dependent manner." Development **134**(15): 2761-2769.

Finn, R. D., A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. Sonnhammer, J. Tate and M. Punta (2014). "Pfam: the protein families database." Nucleic Acids Res **42**(Database issue): D222-230.

Fisher, R. M., J. Z. Shik and J. J. Boomsma (2020). "The evolution of multicellular complexity: the role of relatedness and environmental constraints." Proc Biol Sci **287**(1931): 20192963.

Florey, E., M. A. Cahill and M. Rathmayer (1975). "Excitatory actions of GABA and of acetyl-choline in sea urchin tube feet." Comp Biochem Physiol C Comp Pharmacol **51**(1): 5-12.

Flytzanis, C. N., A. P. McMahon, B. R. Hough-Evans, K. S. Katula, R. J. Britten and E. H. Davidson (1985). "Persistence and integration of cloned DNA in postembryonic sea urchins." Dev Biol **108**(2): 431-442.

Fontaine, J., C. Le Lievre and N. M. Le Douarin (1977). "What is the developmental fate of the neural crest cells which migrate into the pancreas in the avian embryo?" Gen Comp Endocrinol **33**(3): 394-404.

Fujikura, J., K. Hosoda, Y. Kawaguchi, M. Noguchi, H. Iwakura, S. Odori, E. Mori, T. Tomita, M. Hirata, K. Ebihara, H. Masuzaki, A. Fukuda, K. Furuyama, K. Tanigaki, D. Yabe and K. Nakao (2007). "Rbp-j regulates expansion of pancreatic epithelial cells and their differentiation into exocrine cells during mouse development." Dev Dyn **236**(10): 2779-2791.

Fujimoto, K. and K. S. Polonsky (2009). "Pdx1 and other factors that regulate pancreatic beta-cell survival." Diabetes Obes Metab **11 Suppl 4**: 30-37.

Fujita, T., S. Kobayashi and R. Yui (1980). "Paraneuron concept and its current implications." Adv Biochem Psychopharmacol **25**: 321-325.

Fukunaga, K., M. Ohmitsu, E. Miyamoto, T. Sato, M. Sugimura, T. Uchida and Y. Shirasaki (2000). "Inhibition of neuronal nitric oxide synthase activity by 3-[2-[4-(3-chloro-2-methylphenyl)- 1-piperazinyl]ethyl]-5, 6-dimethoxy-1-(4-imidazolylmethyl)-1H-indazole dihydrochloride 3.5 hydrate (DY-9760e), a novel neuroprotective agent, in vitro and in cultured neuroblastoma cells in situ." Biochem Pharmacol **60**(5): 693-699.

Galdino, G. S., C. H. Xavier, R. Almeida, G. Silva, M. A. Fontes, G. Menezes, I. D. Duarte and A. C. Perez (2015). "The Nitric oxide/CGMP/KATP pathway mediates systemic and central antinociception induced by resistance exercise in rats." Int J Neurosci **125**(10): 765-773.

Gao, Y., A. Traulsen and Y. Pichugin (2019). "Interacting cells driving the evolution of multicellular life cycles." PLoS Comput Biol **15**(5): e1006987.

Gao, Y., R. Zhang, S. Dai, X. Zhang, X. Li and C. Bai (2019). "Role of TGF-beta/Smad Pathway in the Transcription of Pancreas-Specific Genes During Beta Cell Differentiation." Front Cell Dev Biol **7**: 351.

Garner, S., I. Zysk, G. Byrne, M. Kramer, D. Moller, V. Taylor and R. D. Burke (2016). "Neurogenesis in sea urchin embryos and the diversity of deuterostome neurogenic mechanisms." Development **143**(2): 286-297.

Geppert, M., Y. A. Ushkaryov, Y. Hata, B. Davletov, A. G. Petrenko and T. C. Sudhof (1992). "Neurexins." Cold Spring Harb Symp Quant Biol **57**: 483-490.

Gilbert, S. and M. Barresi (2018). Developmental biology. New York, NY, Oxford University Press.

Glinka, J., R. Sanchez Claria, V. Ardiles, E. de Santibanes, J. Pekolj, M. de Santibanes and O. Mazza (2019). "The pancreas as a target of metastasis from

renal cell carcinoma: Results of surgical treatment in a single institution." Ann Hepatobiliary Pancreat Surg **23**(3): 240-244.

Gonoi, T., N. Mizuno, N. Inagaki, H. Kuromi, Y. Seino, J. Miyazaki and S. Seino (1994). "Functional neuronal ionotropic glutamate receptors are expressed in the non-neuronal cell line MIN6." J Biol Chem **269**(25): 16989-16992.

Gosmain, Y., L. S. Katz, M. H. Masson, C. Cheyssac, C. Poisson and J. Philippe (2012). "Pax6 is crucial for beta-cell function, insulin biosynthesis, and glucose-induced insulin secretion." Mol Endocrinol **26**(4): 696-709.

Gozes, I., A. Bardea, M. Bechar, O. Pearl, A. Reshef, R. Zamostiano, A. Davidson, S. Rubinraut, E. Giladi, M. Fridkin and D. E. Brenneman (1997). "Neuropeptides and neuronal survival: neuroprotective strategy for Alzheimer's disease." Ann N Y Acad Sci **814**: 161-166.

Gozes, I. and D. E. Brenneman (1993). "Neuropeptides as growth and differentiation factors in general and VIP in particular." J Mol Neurosci **4**(1): 1-9.

Grapin-Botton, A. and D. A. Melton (2000). "Endoderm development: from patterning to organogenesis." Trends Genet **16**(3): 124-130.

Gu, C., G. H. Stein, N. Pan, S. Goebbels, H. Hornberg, K. A. Nave, P. Herrera, P. White, K. H. Kaestner, L. Sussel and J. E. Lee (2010). "Pancreatic beta cells require NeuroD to achieve and maintain functional maturity." Cell Metab **11**(4): 298-310.

Guss, K. A. and C. A. Ettensohn (1997). "Skeletal morphogenesis in the sea urchin embryo: regulation of primary mesenchyme gene expression and skeletal rod growth by ectoderm-derived cues." Development **124**(10): 1899-1908.

Gutierrez, G. D., A. S. Bender, V. Cirulli, T. L. Mastracci, S. M. Kelly, A. Tsirigos, K. H. Kaestner and L. Sussel (2017). "Pancreatic beta cell identity requires continual repression of non-beta cell programs." J Clin Invest **127**(1): 244-259.

Haga, K. K., L. J. Gregory, C. A. Hicks, M. A. Ward, J. S. Beech, P. W. Bath, S. C. Williams and M. J. O'Neill (2003). "The neuronal nitric oxide synthase inhibitor, TRIM, as a neuroprotective agent: effects in models of cerebral ischaemia using histological and magnetic resonance imaging techniques." Brain Res **993**(1-2): 42-53.

Harkey, M. A., H. R. Whiteley and A. H. Whiteley (1992). "Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryo: a three dimensional in situ hybridization analysis." Mech Dev **37**(3): 173-184.

Harlow, P. and M. Nemer (1987). "Coordinate and selective beta-tubulin gene expression associated with cilium formation in sea urchin embryos." Genes Dev **1**(10): 1293-1304.

Hart, A. W., S. Mella, J. Mendrychowski, V. van Heyningen and D. A. Kleinjan (2013). "The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas." PLoS One **8**(1): e54173.

Hartenstein, V. and A. Stollewerk (2015). "The evolution of early neurogenesis." Dev Cell **32**(4): 390-407.

Hay-Schmidt, A. (2000). "The evolution of the serotonergic nervous system." Proc Biol Sci **267**(1448): 1071-1079.

Henquin, J. C. and H. P. Meissner (1984). "Significance of ionic fluxes and changes in membrane potential for stimulus-secretion coupling in pancreatic B-cells." Experientia **40**(10): 1043-1052.

Hinman, V. F. and R. D. Burke (2018). "Embryonic neurogenesis in echinoderms." Wiley Interdiscip Rev Dev Biol **7**(4): e316.

Ho, E. C., K. M. Buckley, C. S. Schrankel, N. W. Schuh, T. Hibino, C. M. Solek, K. Bae, G. Wang and J. P. Rast (2017). "Perturbation of gut bacteria induces a coordinated cellular immune response in the purple sea urchin larva." Immunol Cell Biol **95**(7): 647.

Hodor, P. G. and C. A. Ettensohn (1998). "The dynamics and regulation of mesenchymal cell fusion in the sea urchin embryo." Dev Biol **199**(1): 111-124.

Hoskins R.A., J.W. Carlson, K.H. Wan, S. Park, I. Mendez, S.E. Galle, B.W. Booth, B.D. Pfeiffer, R.A. George, R. Svirskas, M. Krzywinski, J. Schein, M.C. Accardo, E. Damia, G. Messina, M. Méndez-Lago, O.V. de Pablos B, Demakova, E.N. Andreyeva, L.V. Boldyreva, M. Marra, AB Carvalho, P. Dimitri, A. Villasante, I.F. Zhimulev, G.M. Rubin, G.H. Karpen, S.E. Celniker. (2015). “The Release 6 reference sequence of the Drosophila melanogaster genome”. Genome Res. **25**(3):445-58.

Howard-Ashby, M., S. C. Materna, C. T. Brown, L. Chen, R. A. Cameron and E. H. Davidson (2006). "Gene families encoding transcription factors expressed in early development of Strongylocentrotus purpuratus." Dev Biol **300**(1): 90-107.

Hoyle, C. H. (2011). "Evolution of neuronal signalling: transmitters and receptors." Auton Neurosci **165**(1): 28-53.

Hussain, M. A., C. P. Miller and J. F. Habener (2002). "Brn-4 transcription factor expression targeted to the early developing mouse pancreas induces ectopic glucagon gene expression in insulin-producing beta cells." J Biol Chem **277**(18): 16028-16032.

International Human Genome Sequencing Consortium Finishing the euchromatic sequence of the human genome. (2004). Nature. **431**(7011):931–945.

Ishibashi, M., S. L. Ang, K. Shiota, S. Nakanishi, R. Kageyama and F. Guillemot (1995). "Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects." Genes Dev **9**(24): 3136-3148.

Iso, T., L. Kedes and Y. Hamamori (2003). "HES and HERP families: multiple effectors of the Notch signaling pathway." J Cell Physiol **194**(3): 237-255.

Israel, J. W., M. L. Martik, M. Byrne, E. C. Raff, R. A. Raff, D. R. McClay and G. A. Wray (2016). "Comparative Developmental Transcriptomics Reveals Rewiring of a Highly Conserved Gene Regulatory Network during a Major Life History Switch in the Sea Urchin Genus Heliocidaris." PLoS Biol **14**(3): e1002391.

Iturriza, F. C. and J. Thibault (1993). "Immunohistochemical investigation of tyrosine-hydroxylase in the islets of Langerhans of adult mice, rats and guinea pigs." Neuroendocrinology **57**(3): 476-480.

Jekely, G., S. Melzer, I. Beets, I. C. G. Kadow, J. Koene, S. Haddad and L. Holden-Dye (2018). "The long and the short of it - a perspective on peptidergic regulation of circuits and behaviour." Journal of Experimental Biology **221**(3).

Jensen, J. (2004). "Gene regulatory factors in pancreatic development." Dev Dyn **229**(1): 176-200.

Jin, Y., S. Yaguchi, K. Shiba, L. Yamada, J. Yaguchi, D. Shibata, H. Sawada and K. Inaba (2013). "Glutathione transferase theta in apical ciliary tuft regulates mechanical reception and swimming behavior of Sea Urchin Embryos." Cytoskeleton (Hoboken) **70**(8): 453-470.

Juliano, C., S. Z. Swartz and G. Wessel (2014). "Isolating specific embryonic cells of the sea urchin by FACS." Methods Mol Biol **1128**: 187-196.

Juliano, C. E., E. Voronina, C. Stack, M. Aldrich, A. R. Cameron and G. M. Wessel (2006). "Germ line determinants are not localized early in sea urchin development, but do accumulate in the small micromere lineage." Dev Biol **300**(1): 406-415.

Juliano, C. E., M. Yajima and G. M. Wessel (2010). "Nanos functions to maintain the fate of the small micromere lineage in the sea urchin embryo." Dev Biol **337**(2): 220-232.

Jung, M., D. Wells, J. Rusch, S. Ahmad, J. Marchini, S. R. Myers and D. F. Conrad (2019). "Unified single-cell analysis of testis gene regulation and pathology in five mouse strains." Elife **8**.

Kashima, Y., Y. Sakamoto, K. Kaneko, M. Seki, Y. Suzuki and A. Suzuki (2020). "Single-cell sequencing techniques from individual to multiomics analyses." Exp Mol Med **52**(9): 1419-1427.

Kashiwagi, S., M. Kajimura, Y. Yoshimura and M. Suematsu (2002). "Nonendothelial source of nitric oxide in arterioles but not in venules: alternative source revealed in vivo by diaminofluorescein microfluorography." Circ Res **91**(12): e55-64.

Katow, H., K. Abe, T. Katow, A. Zamani and H. Abe (2013). "Development of the GABA-ergic signaling system and its role in larval swimming in sea urchin." J Exp Biol **216**(Pt 9): 1704-1716.

Katow, H., T. Katow, K. Abe, S. Ooka, M. Kiyomoto and G. Hamanaka (2014). "Mesomere-derived glutamate decarboxylase-expressing blastocoelar mesenchyme cells of sea urchin larvae." Biol Open **3**(1): 94-102.

Katow, H., T. Suyemitsu, S. Ooka, J. Yaguchi, T. Jin-Nai, I. Kuwahara, T. Katow, S. Yaguchi and H. Abe (2010). "Development of a dopaminergic system in sea urchin embryos and larvae." J Exp Biol **213**(Pt 16): 2808-2819.

Katow, H., S. Yaguchi and K. Kyozuka (2007). "Serotonin stimulates [Ca2+]i elevation in ciliary ectodermal cells of echinoplutei through a serotonin receptor cell network in the blastocoel." J Exp Biol **210**(Pt 3): 403-412.

Katow, H., H. Yoshida and M. Kiyomoto (2020). "Initial report of gamma-aminobutyric acidergic locomotion regulatory system and its 3-mercaptopropionic acid-sensitivity in metamorphic juvenile of sea urchin, Hemicentrotus pulcherrimus." Sci Rep **10**(1): 778.

Kawaguchi, Y., B. Cooper, M. Gannon, M. Ray, R. J. MacDonald and C. V. Wright (2002). "The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors." Nat Genet **32**(1): 128-134.

Kim, W. Y., B. Fritzsch, A. Serls, L. A. Bakel, E. J. Huang, L. F. Reichardt, D. S. Barth and J. E. Lee (2001). "NeuroD-null mice are deaf due to a severe loss of the inner ear sensory neurons during development." Development **128**(3): 417-426.

Kinjo, S., M. Kiyomoto, T. Yamamoto, K. Ikeo and S. Yaguchi (2018). "HpBase: A genome database of a sea urchin, Hemicentrotus pulcherrimus." Dev Growth Differ **60**(3): 174-182.
Kipryushina, Y. O. and K. V. Yakovlev (2020). "Maternal control of early patterning in sea urchin embryos." Differentiation **113**: 28-37.

Kirchhausen, T., D. Owen and S. C. Harrison (2014). "Molecular structure, function, and dynamics of clathrin-mediated membrane traffic." Cold Spring Harb Perspect Biol **6**(5): a016725.

Kojima, H., Y. Urano, K. Kikuchi, T. Higuchi, Y. Hirata and T. Nagano (1999). "Fluorescent Indicators for Imaging Nitric Oxide Production." Angew Chem Int Ed Engl **38**(21): 3209-3212.

Kolluru, G. K., K. P. Tamilarasan, A. S. Rajkumar, S. Geetha Priya, M. Rajaram, N. K. Saleem, S. Majumder, B. M. Jaffar Ali, G. Illavazagan and S. Chatterjee (2008). "Nitric oxide/cGMP protects endothelial cells from hypoxia-mediated leakiness." Eur J Cell Biol **87**(3): 147-161.

Krouk, G., J. Lingeman, A. M. Colon, G. Coruzzi and D. Shasha (2013). "Gene regulatory networks in plants: learning causality from time and perturbation." Genome Biol **14**(6): 123.

Kudtarkar, P. and R. A. Cameron (2017). "Echinobase: an expanding resource for echinoderm genomic information." Database (Oxford) **2017**.

Langley, O. K., M. C. Aletsee-Ufrecht, N. J. Grant and M. Gratzl (1989). "Expression of the neural cell adhesion molecule NCAM in endocrine cells." J Histochem Cytochem **37**(6): 781-791.

Lapraz, F., L. Besnardeau and T. Lepage (2009). "Patterning of the dorsal-ventral axis in echinoderms: insights into the evolution of the BMP-chordin signaling network." PLoS Biol **7**(11): e1000248.

Lapraz, F., E. Haillot and T. Lepage (2015). "A deuterostome origin of the Spemann organiser suggested by Nodal and ADMPs functions in Echinoderms." Nat Commun **6**: 8927.

Lasky, J. L. and H. Wu (2005). "Notch signaling, brain development, and human disease." Pediatr Res **57**(5 Pt 2): 104R-109R.

Lecroisey, C., Y. Le Petillon, H. Escriva, E. Lammert and V. Laudet (2015). "Identification, evolution and expression of an insulin-like peptide in the cephalochordate Branchiostoma lanceolatum." PLoS One **10**(3): e0119461.

Lemak, M. S. (2012). "[Endocannabinoid signalling in the central nervous system of vertebrates and invertebrates]." Zh Vyssh Nerv Deiat Im I P Pavlova **62**(5): 531-543.

Levine, M. and E. H. Davidson (2005). "Gene regulatory networks for development." Proc Natl Acad Sci U S A **102**(14): 4936-4942.

Li, H., S. Arber, T. M. Jessell and H. Edlund (1999). "Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlxb9." Nat Genet **23**(1): 67-70.

Lin, C. Y., N. Oulhen, G. Wessel and Y. H. Su (2019). "CRISPR/Cas9-mediated genome editing in sea urchins." Methods Cell Biol **151**: 305-321.

Lin, C. Y. and Y. H. Su (2016). "Genome editing in sea urchin embryos by using a CRISPR/Cas9 system." Dev Biol **409**(2): 420-428.

Liu, D., A. Awazu, T. Sakuma, T. Yamamoto and N. Sakamoto (2019). "Establishment of knockout adult sea urchins by using a CRISPR-Cas9 system." Dev Growth Differ **61**(6): 378-388.

Logan, C. Y., J. R. Miller, M. J. Ferkowicz and D. R. McClay (1999). "Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo." Development **126**(2): 345-357.

Longabaugh, W. J. (2012). "BioTapestry: a tool to visualize the dynamic properties of gene regulatory networks." Methods Mol Biol **786**: 359-394.

Loosli, F., R. W. Koster, M. Carl, A. Krone and J. Wittbrodt (1998). "Six3, a medaka homologue of the Drosophila homeobox gene sine oculis is expressed in the anterior embryonic shield and the developing eye." Mech Dev **74**(1-2): 159-164.

Lowe E.K., C. Cuomo, M.I. Arnone. (2016). "A Differential Transcriptomic Approach to Compare Target Genes of Homologous Transcription Factors in Echinoderm Species". In: Rogato A., Zazzu V., Guarracino M. (eds) Dynamics of Mathematical Models in Biology. Springer, Cham.

Luo, Q., Q. Fu, X. Zhang, H. Zhang and T. Qin (2020). "Application of Single-Cell RNA Sequencing in Pancreatic Cancer and the Endocrine Pancreas." Adv Exp Med Biol **1255**: 143-152.

Luo, Y. J. and Y. H. Su (2012). "Opposing nodal and BMP signals regulate left-right asymmetry in the sea urchin larva." PLoS Biol **10**(10): e1001402.

Macfarlane, W. M., C. M. McKinnon, Z. A. Felton-Edkins, H. Cragg, R. F. James and K. Docherty (1999). "Glucose stimulates translocation of the homeodomain transcription factor PDX1 from the cytoplasm to the nucleus in pancreatic beta-cells." J Biol Chem **274**(2): 1011-1016.

Machon, O., M. Backman, O. Machonova, Z. Kozmik, T. Vacik, L. Andersen and S. Krauss (2007). "A dynamic gradient of Wnt signaling controls initiation of neurogenesis in the mammalian cortex and cellular specification in the hippocampus." Dev Biol **311**(1): 223-237.

Maechler, P. and C. B. Wollheim (1999). "Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis." Nature **402**(6762): 685-689.

Magri, M. S., S. Jimenez-Gancedo, S. Bertrand, A. Madgwick, H. Escriva, P. Lemaire and J. L. Gomez-Skarmeta (2019). "Assaying Chromatin Accessibility Using ATAC-Seq in Invertebrate Chordate Embryos." Front Cell Dev Biol **7**: 372.

Marinkovic, M., J. Berger and G. Jekely (2020). "Neuronal coordination of motile cilia in locomotion and feeding." Philos Trans R Soc Lond B Biol Sci **375**(1792): 20190165.

Marlow, H., M. A. Tosches, R. Tomer, P. R. Steinmetz, A. Lauri, T. Larsson and D. Arendt (2014). "Larval body patterning and apical organs are conserved in animal evolution." BMC Biol **12**: 7.

Martella, A., R. M. Sepe, C. Silvestri, J. Zang, G. Fasano, O. Carnevali, P. De Girolamo, S. C. Neuhauss, P. Sordino and V. Di Marzo (2016). "Important role of endocannabinoid signaling in the development of functional vision and locomotion in zebrafish." FASEB J **30**(12): 4275-4288.

Martik, M. L., D. C. Lyons and D. R. McClay (2016). "Developmental gene regulatory networks in sea urchins and what we can learn from them." F1000Res **5**.

Martik, M. L. and D. R. McClay (2015). "Deployment of a retinal determination gene network drives directed cell migration in the sea urchin embryo." Elife **4**.

Martinelli, P., M. Canamero, N. del Pozo, F. Madriles, A. Zapata and F. X. Real (2013). "Gata6 is required for complete acinar differentiation and maintenance of the exocrine pancreas in adult mice." Gut **62**(10): 1481-1488.

Martinez-Bartolome, M. and R. C. Range (2019). "A biphasic role of non-canonical Wnt16 signaling during early anterior-posterior patterning and morphogenesis of the sea urchin embryo." Development **146**(24).

Martinez-Ruiz, A., S. Cadenas and S. Lamas (2011). "Nitric oxide signaling: classical, less classical, and nonclassical mechanisms." Free Radic Biol Med **51**(1): 17-29.

Martinez, A. (1995). "Nitric oxide synthase in invertebrates." Histochem J **27**(10): 770-776.
Marty-Santos, L. and O. Cleaver (2016). "Pdx1 regulates pancreas tubulogenesis and E-cadherin expression." Development **143**(6): 1056.

Massri, A. J., L. Greenstreet, A. Afanassiev, A. B. Escobar, G. M. Wray, G. Schiebinger and D. R. McClay (2020). "Developmental Single-cell transcriptomics in the Lytechinus variegatus Sea Urchin Embryo." bioRxiv: 2020.2011.2012.380675.

Materna, S. C. (2017). "Using Morpholinos to Probe Gene Networks in Sea Urchin." Methods Mol Biol **1565**: 87-104.

Materna, S. C., A. Ransick, E. Li and E. H. Davidson (2013). "Diversification of oral and aboral mesodermal regulatory states in pregastrular sea urchin embryos." Dev Biol **375**(1): 92-104.

McClay, D. R. (1986). "Embryo dissociation, cell isolation, and cell reassociation." Methods Cell Biol **27**: 309-323.

McClay, D. R. (2004). "Methods for embryo dissociation and analysis of cell adhesion." Methods Cell Biol **74**: 311-329.

McClay, D. R., E. Miranda and S. L. Feinberg (2018). "Neurogenesis in the sea urchin embryo is initiated uniquely in three domains." Development **145**(21).

McIntyre, D. C., D. C. Lyons, M. Martik and D. R. McClay (2014). "Branching out: origins of the sea urchin larval skeleton in development and evolution." Genesis **52**(3): 173-185.

McQueen, E. and M. Rebeiz (2020). "On the specificity of gene regulatory networks: How does network co-option affect subsequent evolution?" Curr Top Dev Biol **139**: 375-405.

Medeiros Filho, F., A. P. B. do Nascimento, M. T. Dos Santos, A. P. D. Carvalho-Assef and F. A. B. da Silva (2019). "Gene regulatory network inference and analysis of multidrug-resistant Pseudomonas aeruginosa." Mem Inst Oswaldo Cruz **114**: e190105.

Mellott, D. O., J. Thisdelle and R. D. Burke (2017). "Notch signaling patterns neurogenic ectoderm and regulates the asymmetric division of neural progenitors in sea urchin embryos." Development **144**(19): 3602-3611.

Meyers, E. A. and J. A. Kessler (2017). "TGF-beta Family Signaling in Neural and Neuronal Differentiation, Development, and Function." Cold Spring Harb Perspect Biol **9**(8).

Mitchell, R. K., M. S. Nguyen-Tu, P. Chabosseau, R. M. Callingham, T. J. Pullen, R. Cheung, I. Leclerc, D. J. Hodson and G. A. Rutter (2017). "The transcription factor Pax6 is required for pancreatic beta cell identity, glucose-regulated ATP synthesis, and Ca(2+) dynamics in adult mice." J Biol Chem **292**(21): 8892-8906.

Mizusawa, N., T. Hasegawa, I. Ohigashi, C. Tanaka-Kosugi, N. Harada, M. Itakura and K. Yoshimoto (2004). "Differentiation phenotypes of pancreatic islet beta- and alpha-cells are closely related with homeotic genes and a group of differentially expressed genes." Gene **331**: 53-63.

Muraro, M. J., G. Dharmadhikari, D. Grun, N. Groen, T. Dielen, E. Jansen, L. van Gurp, M. A. Engelse, F. Carlotti, E. J. de Koning and A. van Oudenaarden (2016). "A Single-Cell Transcriptome Atlas of the Human Pancreas." Cell Syst **3**(4): 385-394 e383.

Natalie J. Wood, T. M., Elizabeth Ward, Matthew L. Rowe, Margherita Perillo, M. Ina Arnone , Maurice R. Elphick and Paola Oliveri (2018). "Neuropeptidergic systems in pluteus larvae of the sea urchin Strongylocentrotus purpuratus: neurochemical complexity in a “simple” nervous system." Fronties in Endocrinology.

Navale, A. M. and A. N. Paranjape (2016). "Glucose transporters: physiological and pathological roles." Biophys Rev **8**(1): 5-9.

Nestorowa, S., F. K. Hamey, B. Pijuan Sala, E. Diamanti, M. Shepherd, E. Laurenti, N. K. Wilson, D. G. Kent and B. Gottgens (2016). "A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation." Blood **128**(8): e20-31.

Nishimaru, H., C. E. Restrepo, J. Ryge, Y. Yanagawa and O. Kiehn (2005). "Mammalian motor neurons corelease glutamate and acetylcholine at central synapses." Proc Natl Acad Sci U S A **102**(14): 5245-5249.

O'Sullivan, S. E. (2007). "Cannabinoids go nuclear: evidence for activation of peroxisome proliferator-activated receptors." Br J Pharmacol **152**(5): 576-582.

O'Sullivan, S. E. (2016). "An update on PPAR activation by cannabinoids." Br J Pharmacol **173**(12): 1899-1910.

Olinski, R. P., C. Dahlberg, M. Thorndyke and F. Hallbook (2006). "Three insulin-relaxin-like genes in Ciona intestinalis." Peptides **27**(11): 2535-2546.

Oliver-Krasinski, J. M. and D. A. Stoffers (2008). "On the origin of the beta cell." Genes Dev **22**(15): 1998-2021.

Oulhen, N. and G. M. Wessel (2016). "Albinism as a visual, in vivo guide for CRISPR/Cas9 functionality in the sea urchin embryo." Mol Reprod Dev **83**(12): 1046-1047.

Palumbo, A. (2005). "Nitric oxide in marine invertebrates: a comparative perspective." Comp Biochem Physiol A Mol Integr Physiol **142**(2): 241-248.

Pan, F. C., M. Brissova, A. C. Powers, S. Pfaff and C. V. Wright (2015). "Inactivating the permanent neonatal diabetes gene Mnx1 switches insulin-producing beta-cells to a delta-like fate and reveals a facultative proliferative capacity in aged beta-cells." Development **142**(21): 3637-3648.

Pandey, G., E. Makhija, N. George, B. Chakravarti, M. M. Godbole, C. M. Ecelbarger and S. Tiwari (2015). "Insulin regulates nitric oxide production in the kidney collecting duct cells." J Biol Chem **290**(9): 5582-5591.

Park, D., O. T. Shafer, S. P. Shepherd, H. Suh, J. S. Trigg and P. H. Taghert (2008). "The Drosophila basic helix-loop-helix protein DIMMED directly activates PHM, a gene encoding a neuropeptide-amidating enzyme." Mol Cell Biol **28**(1): 410-421.

Park, H. S., S. H. Huh, M. S. Kim, S. H. Lee and E. J. Choi (2000). "Nitric oxide negatively regulates c-Jun N-terminal kinase/stress-activated protein kinase by means of S-nitrosylation." Proc Natl Acad Sci U S A **97**(26): 14382-14387.

Pearl, E. J., C. K. Bilogan, S. Mukhi, D. D. Brown and M. E. Horb (2009). "Xenopus pancreas development." Dev Dyn **238**(6): 1271-1286.

Pearse, A. G. and J. M. Polak (1971). "Neural crest origin of the endocrine polypeptide (APUD) cells of the gastrointestinal tract and pancreas." Gut **12**(10): 783-788.

Pehrson, J. R. and L. H. Cohen (1986). "The fate of the small micromeres in sea urchin development." Dev Biol **113**(2): 522-526.

Pei, D. S., Y. J. Song, H. M. Yu, W. W. Hu, Y. Du and G. Y. Zhang (2008). "Exogenous nitric oxide negatively regulates c-Jun N-terminal kinase activation via inhibiting endogenous NO-induced S-nitrosylation during cerebral ischemia and reperfusion in rat hippocampus." J Neurochem **106**(4): 1952-1963.

Perelis, M., B. Marcheva, K. M. Ramsey, M. J. Schipma, A. L. Hutchison, A. Taguchi, C. B. Peek, H. Hong, W. Huang, C. Omura, A. L. Allred, C. A. Bradfield, A. R. Dinner, G. D. Barish and J. Bass (2015). "Pancreatic beta cell enhancers regulate rhythmic transcription of genes controlling insulin secretion." Science **350**(6261): aac4250.

Perillo, M. (2013). Evolutionary Origin and Diversification of Pancreatic Cell Types During Sea Urchin Gut Morphogenesis, Open University.

Perillo, M. and M. I. Arnone (2014). "Characterization of insulin-like peptides (ILPs) in the sea urchin Strongylocentrotus purpuratus: insights on the evolution of the insulin family." Gen Comp Endocrinol **205**: 68-79.

Perillo, M., N. Oulhen, S. Foster, M. Spurrell, C. Calestani and G. Wessel (2020). "Regulation of dynamic pigment cell states at single-cell resolution." Elife **9**.

Perillo, M., P. Paganos, T. Mattiello, M. Cocurullo, P. Oliveri and M. I. Arnone (2018). "New Neuronal Subtypes With a "Pre-Pancreatic" Signature in the Sea Urchin Stongylocentrotus purpuratus." Front Endocrinol (Lausanne) **9**: 650.

Perillo, M., P. Paganos, M. Spurrell, M. I. Arnone and G. M. Wessel (2021). "Methodology for Whole Mount and Fluorescent RNA In Situ Hybridization in Echinoderms: Single, Double, and Beyond." Methods Mol Biol **2219**: 195-216.

Perillo, M., Y. J. Wang, S. D. Leach and M. I. Arnone (2016). "A pancreatic exocrine-like cell regulatory circuit operating in the upper stomach of the sea urchin Strongylocentrotus purpuratus larva." BMC Evol Biol **16**(1): 117.

Piccand, J., P. Strasser, D. J. Hodson, A. Meunier, T. Ye, C. Keime, M. C. Birling, G. A. Rutter and G. Gradwohl (2014). "Rfx6 maintains the functional identity of adult pancreatic beta cells." Cell Rep **9**(6): 2219-2232.

Pictet, R. L., L. B. Rall, P. Phelps and W. J. Rutter (1976). "The neural crest and the origin of the insulin-producing and other gastrointestinal hormone-producing cells." Science **191**(4223): 191-192.

Pin, C. L., J. M. Rukstalis, C. Johnson and S. F. Konieczny (2001). "The bHLH transcription factor Mist1 is required to maintain exocrine pancreas cell organization and acinar cell identity." J Cell Biol **155**(4): 519-530.

Posnien, N., N. D. Koniszewski, H. J. Hein and G. Bucher (2011). "Candidate gene screen in the red flour beetle Tribolium reveals six3 as ancient regulator of anterior median head and central complex development." PLoS Genet **7**(12): e1002416.

Qi, F., S. Qian, S. Zhang and Z. Zhang (2020). "Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses." Biochem Biophys Res Commun **526**(1): 135-140.

Qu, X., S. Afelik, J. N. Jensen, M. A. Bukys, S. Kobberup, M. Schmerr, F. Xiao, P. Nyeng, M. Veronica Albertoni, A. Grapin-Botton and J. Jensen (2013). "Notch-mediated post-translational control of Ngn3 protein stability regulates pancreatic patterning and cell fate commitment." Dev Biol **376**(1): 1-12.

Rafiq, K., M. S. Cheers and C. A. Ettensohn (2012). "The genomic regulatory control of skeletal morphogenesis in the sea urchin." Development **139**(3): 579-590.
Rafiq, K., T. Shashikant, C. J. McManus and C. A. Ettensohn (2014). "Genome-wide analysis of the skeletogenic gene regulatory network of sea urchins." Development **141**(4): 950-961.

Ramond, C., B. S. Beydag-Tasoz, A. Azad, M. van de Bunt, M. B. K. Petersen, N. L. Beer, N. Glaser, C. Berthault, A. L. Gloyn, M. Hansson, M. I. McCarthy, C. Honore, A. Grapin-Botton and R. Scharfmann (2018). "Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling." Development **145**(16).

Range, R. (2014). "Specification and positioning of the anterior neuroectoderm in deuterostome embryos." Genesis **52**(3): 222-234.

Range, R. and T. Lepage (2011). "Maternal Oct1/2 is required for Nodal and Vg1/Univin expression during dorsal-ventral axis specification in the sea urchin embryo." Dev Biol **357**(2): 440-449.

Range, R. C., R. C. Angerer and L. M. Angerer (2013). "Integration of canonical and noncanonical Wnt signaling pathways patterns the neuroectoderm along the anterior-posterior axis of sea urchin embryos." PLoS Biol **11**(1): e1001467.

Range, R. C. and Z. Wei (2017). "Correction: An anterior signaling center patterns and sizes the anterior neuroectoderm of the sea urchin embryo." Development **144**(8): 1579.

Ratcliff, W. C., R. F. Denison, M. Borrello and M. Travisano (2012). "Experimental evolution of multicellularity." Proc Natl Acad Sci U S A **109**(5): 1595-1600.

Ray, J. D., K. B. Kener, B. F. Bitner, B. J. Wright, M. S. Ballard, E. J. Barrett, J. T. Hill, L. G. Moss and J. S. Tessem (2016). "Nkx6.1-mediated insulin secretion and beta-cell proliferation is dependent on upregulation of c-Fos." FEBS Lett **590**(12): 1791-1803.

Reetz, A., M. Solimena, M. Matteoli, F. Folli, K. Takei and P. De Camilli (1991). "GABA and pancreatic beta-cells: colocalization of glutamic acid decarboxylase (GAD) and GABA with synaptic-like microvesicles suggests their role in GABA storage and secretion." EMBO J **10**(5): 1275-1284.

Reinecke, M. and C. Collet (1998). "The phylogeny of the insulin-like growth factors." Int Rev Cytol **183**: 1-94.

Rhinehart, K. L. and T. L. Pallone (2001). "Nitric oxide generation by isolated descending vasa recta." Am J Physiol Heart Circ Physiol **281**(1): H316-324.

Rindi, G., A. B. Leiter, A. S. Kopin, C. Bordi and E. Solcia (2004). "The "normal" endocrine cell of the gut: changing concepts and new evidences." Ann N Y Acad Sci **1014**: 1-12.

Rindi, G., A. Torsello, V. Locatelli and E. Solcia (2004). "Ghrelin expression and actions: a novel peptide for an old cell type of the diffuse endocrine system." Exp Biol Med (Maywood) **229**(10): 1007-1016.

Rivero, A. (2006). "Nitric oxide: an antiparasitic molecule of invertebrates." Trends Parasitol **22**(5): 219-225.

Rizzo, F., M. Fernandez-Serra, P. Squarzoni, A. Archimandritis and M. I. Arnone (2006). "Identification and developmental expression of the ets gene family in the sea urchin (Strongylocentrotus purpuratus)." Dev Biol **300**(1): 35-48.

Rodriguez-Diaz, R., R. Dando, M. C. Jacques-Silva, A. Fachado, J. Molina, M. H. Abdulreda, C. Ricordi, S. D. Roper, P. O. Berggren and A. Caicedo (2011). "Alpha cells secrete acetylcholine as a non-neuronal paracrine signal priming beta cell function in humans." Nat Med **17**(7): 888-892.

Rowe, M. L. and M. R. Elphick (2012). "The neuropeptide transcriptome of a model echinoderm, the sea urchin Strongylocentrotus purpuratus." Gen Comp Endocrinol **179**(3): 331-344.

Ruppert, E. E., R. S. Fox and R. D. Barnes. (2004). "Invertebrate zoology : a functional evolutionary approach." from http://catalog.hathitrust.org/api/volumes/oclc/53021401.html.

Salzet, M. and G. B. Stefano (2002). "The endocannabinoid system in invertebrates." Prostaglandins Leukot Essent Fatty Acids **66**(2-3): 353-361.

Saudemont, A., E. Haillot, F. Mekpoh, N. Bessodes, M. Quirin, F. Lapraz, V. Duboc, E. Rottinger, R. Range, A. Oisel, L. Besnardeau, P. Wincker and T. Lepage (2010). "Ancestral regulatory circuits governing ectoderm patterning downstream of Nodal and BMP2/4 revealed by gene regulatory network analysis in an echinoderm." PLoS Genet **6**(12): e1001259.

Saunders, L. R. and D. R. McClay (2014). "Sub-circuits of a gene regulatory network control a developmental epithelial-mesenchymal transition." Development **141**(7): 1503-1513.

Schuurmans, C. and F. Guillemot (2002). "Molecular mechanisms underlying cell fate specification in the developing telencephalon." Curr Opin Neurobiol **12**(1): 26-34.

Schwitzgebel, V. M., D. W. Scheel, J. R. Conners, J. Kalamaras, J. E. Lee, D. J. Anderson, L. Sussel, J. D. Johnson and M. S. German (2000). "Expression of neurogenin3 reveals an islet cell precursor population in the pancreas." Development **127**(16): 3533-3542.

Sea Urchin Genome Sequencing, C., E. Sodergren, G. M. Weinstock, E. H. Davidson, R. A. Cameron, R. A. Gibbs, R. C. Angerer, L. M. Angerer, M. I. Arnone, D. R. Burgess, R. D. Burke, J. A. Coffman, M. Dean, M. R. Elphick, C. A. Ettensohn, K. R. Foltz, A. Hamdoun, R. O. Hynes, W. H. Klein, W. Marzluff, D. R. McClay, R. L. Morris, A. Mushegian, J. P. Rast, L. C. Smith, M. C. Thorndyke, V. D. Vacquier, G. M. Wessel, G. Wray, L. Zhang, C. G. Elsik, O. Ermolaeva, W. Hlavina, G. Hofmann, P. Kitts, M. J. Landrum, A. J. Mackey, D. Maglott, G. Panopoulou, A. J. Poustka, K. Pruitt, V. Sapojnikov, X. Song, A. Souvorov, V. Solovyev, Z. Wei, C. A. Whittaker, K. Worley, K. J. Durbin, Y. Shen, O. Fedrigo, D. Garfield, R. Haygood, A. Primus, R. Satija, T. Severson, M. L. Gonzalez-Garay, A. R. Jackson, A. Milosavljevic, M. Tong, C. E. Killian, B. T. Livingston, F. H. Wilt, N. Adams, R. Belle, S. Carbonneau, R. Cheung, P. Cormier, B. Cosson, J. Croce, A. Fernandez-Guerra, A. M. Geneviere, M. Goel, H. Kelkar, J. Morales, O. Mulner-Lorillon, A. J. Robertson, J. V. Goldstone, B. Cole, D. Epel, B. Gold, M. E. Hahn, M. Howard-Ashby, M. Scally, J. J. Stegeman, E. L. Allgood, J. Cool, K. M. Judkins, S. S. McCafferty, A. M. Musante, R. A. Obar, A. P. Rawson, B. J. Rossetti, I. R. Gibbons, M. P. Hoffman, A. Leone, S. Istrail, S. C. Materna, M. P. Samanta, V. Stolc, W. Tongprasit, Q. Tu, K. F. Bergeron, B. P. Brandhorst, J. Whittle, K. Berney, D. J. Bottjer, C. Calestani, K. Peterson, E. Chow, Q. A. Yuan, E. Elhaik, D. Graur, J. T. Reese, I. Bosdet, S. Heesun, M. A. Marra, J. Schein, M. K. Anderson, V. Brockton, K. M. Buckley, A. H. Cohen, S. D. Fugmann, T. Hibino, M. Loza-Coll, A. J. Majeske, C. Messier, S. V. Nair, Z. Pancer, D. P. Terwilliger, C. Agca, E. Arboleda, N. Chen, A. M. Churcher, F. Hallbook, G. W. Humphrey, M. M. Idris, T. Kiyama, S. Liang, D. Mellott, X. Mu, G. Murray, R. P. Olinski, F. Raible, M. Rowe, J. S. Taylor, K. Tessmar-Raible, D. Wang, K. H. Wilson, S. Yaguchi, T. Gaasterland, B. E. Galindo, H. J. Gunaratne, C. Juliano, M. Kinukawa, G. W. Moy, A. T. Neill, M. Nomura, M. Raisch, A. Reade, M. M. Roux, J. L. Song, Y. H. Su, I. K. Townley, E. Voronina, J. L. Wong, G. Amore, M. Branno, E. R. Brown, V. Cavalieri, V. Duboc, L. Duloquin, C. Flytzanis, C. Gache, F. Lapraz, T. Lepage, A. Locascio, P. Martinez, G. Matassi, V. Matranga, R. Range, F. Rizzo, E. Rottinger, W. Beane, C. Bradham, C. Byrum, T. Glenn, S. Hussain, G.

Manning, E. Miranda, R. Thomason, K. Walton, A. Wikramanayke, S. Y. Wu, R. Xu, C. T. Brown, L. Chen, R. F. Gray, P. Y. Lee, J. Nam, P. Oliveri, J. Smith, D. Muzny, S. Bell, J. Chacko, A. Cree, S. Curry, C. Davis, H. Dinh, S. Dugan-Rocha, J. Fowler, R. Gill, C. Hamilton, J. Hernandez, S. Hines, J. Hume, L. Jackson, A. Jolivet, C. Kovar, S. Lee, L. Lewis, G. Miner, M. Morgan, L. V. Nazareth, G. Okwuonu, D. Parker, L. L. Pu, R. Thorn and R. Wright (2006). "The genome of the sea urchin Strongylocentrotus purpuratus." Science **314**(5801): 941-952.

Sebe-Pedros, A., B. Saudemont, E. Chomsky, F. Plessier, M. P. Mailhe, J. Renno, Y. Loe-Mie, A. Lifshitz, Z. Mukamel, S. Schmutz, S. Novault, P. R. H. Steinmetz, F. Spitz, A. Tanay and H. Marlow (2018). "Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq." Cell **173**(6): 1520-1534 e1520.

Sethi, A. J., R. C. Angerer and L. M. Angerer (2009). "Gene regulatory network interactions in sea urchin endomesoderm induction." PLoS Biol **7**(2): e1000029.

Severo, M. S., J. J. M. Landry, R. L. Lindquist, C. Goosmann, V. Brinkmann, P. Collier, A. E. Hauser, V. Benes, J. Henriksson, S. A. Teichmann and E. A. Levashina (2018). "Unbiased classification of mosquito blood cells by single-cell genomics and high-content imaging." Proc Natl Acad Sci U S A **115**(32): E7568-E7577.

Sharma, A., M. Moore, E. Marcora, J. E. Lee, Y. Qiu, S. Samaras and R. Stein (1999). "The NeuroD1/BETA2 sequences essential for insulin gene transcription colocalize with those necessary for neurogenesis and p300/CREB binding protein binding." Mol Cell Biol **19**(1): 704-713.

Sharma, S., W. Wang and A. Stolfi (2019). "Single-cell transcriptome profiling of the Ciona larval brain." Dev Biol **448**(2): 226-236.

Sharma, T. and C. A. Ettensohn (2011). "Regulative deployment of the skeletogenic gene regulatory network during sea urchin development." Development **138**(12): 2581-2590.

Shashikant, T., J. M. Khor and C. A. Ettensohn (2018). "From genome to anatomy: The architecture and evolution of the skeletogenic gene regulatory network of sea urchins and other echinoderms." Genesis **56**(10): e23253.

Shekhar, K. and V. Menon (2019). "Identification of Cell Types from Single-Cell Transcriptomic Data." Methods Mol Biol **1935**: 45-77.

Shen-Orr, S. S., Y. Pilpel and C. P. Hunter (2010). "Composition and regulation of maternal and zygotic transcriptomes reflects species-specific reproductive mode." Genome Biol **11**(6): R58.

Simon-Areces, J., G. Membrive, C. Garcia-Fernandez, L. M. Garcia-Segura and M. A. Arevalo (2010). "Neurogenin 3 cellular and subcellular localization in the developing and adult hippocampus." J Comp Neurol **518**(10): 1814-1824.

Slack, J. M. (1995). "Developmental biology of the pancreas." Development **121**(6): 1569-1580.

Slota, L. A. and D. R. McClay (2018). "Identification of neural transcription factors required for the differentiation of three neuronal subtypes in the sea urchin embryo." Dev Biol **435**(2): 138-149.

Smith, M. M., L. Cruz Smith, R. A. Cameron and L. A. Urry (2008). "The larval stages of the sea urchin, Strongylocentrotus purpuratus." J Morphol **269**(6): 713-733.

Smith, S. B., H. Q. Qu, N. Taleb, N. Y. Kishimoto, D. W. Scheel, Y. Lu, A. M. Patch, R. Grabs, J. Wang, F. C. Lynn, T. Miyatsuka, J. Mitchell, R. Seerke, J. Desir, S. Vanden Eijnden, M. Abramowicz, N. Kacet, J. Weill, M. E. Renard, M. Gentile, I. Hansen, K. Dewar, A. T. Hattersley, R. Wang, M. E. Wilson, J. D. Johnson, C. Polychronakos and M. S. German (2010). "Rfx6 directs islet formation and insulin production in mice and humans." Nature **463**(7282): 775-780.

Soliman, S. (1984). "Pharmacological control of ciliary activity in the young sea urchin larva. Studies on the role of Ca2+ and cyclic nucleotides." Comp Biochem Physiol C Comp Pharmacol Toxicol **78**(1): 183-191.

Salomon R, D. Kaczorowski, F. Valdes-Mora, R.E. Nordon, A. Neild, N. Farbehi, N. Bartonicek, D. Gallego-Ortega. (2019). "Droplet-based single cell RNAseq tools: a practical guide". Lab Chip. 14;19(10):1706-1727.

Song, J., Y. Xu, X. Hu, B. Choi and Q. Tong (2010). "Brain expression of Cre recombinase driven by pancreas-specific promoters." Genesis **48**(11): 628-634.

Squires, L. N., S. S. Rubakhin, A. A. Wadhams, K. N. Talbot, H. Nakano, L. L. Moroz and J. V. Sweedler (2010). "Serotonin and its metabolism in basal deuterostomes: insights from Strongylocentrotus purpuratus and Xenoturbella bocki." J Exp Biol **213**(Pt 15): 2647-2654.

St Laurent, C. D., T. C. Moon and A. D. Befus (2015). "Measurement of nitric oxide in mast cells with the fluorescent indicator DAF-FM diacetate." Methods Mol Biol **1220**: 339-345.

Stamler, J. S., D. I. Simon, J. A. Osborne, M. E. Mullins, O. Jaraki, T. Michel, D. J. Singel and J. Loscalzo (1992). "S-nitrosylation of proteins with nitric oxide: synthesis and characterization of biologically active compounds." Proc Natl Acad Sci U S A **89**(1): 444-448.

Steinmetz, P. R., R. Urbach, N. Posnien, J. Eriksson, R. P. Kostyuchenko, C. Brena, K. Guy, M. Akam, G. Bucher and D. Arendt (2010). "Six3 demarcates the anterior-most developing brain region in bilaterian animals." Evodevo **1**(1): 14.

Steinmetz, P. R. H., A. Aman, J. E. M. Kraus and U. Technau (2017). "Gut-like ectodermal tissue in a sea anemone challenges germ layer homology." Nat Ecol Evol **1**(10): 1535-1542.

Strand, F. L., K. J. Rose, L. A. Zuccarelli, J. Kume, S. E. Alves, F. J. Antonawich and L. Y. Garrett (1991). "Neuropeptide hormones as neurotrophic factors." Physiol Rev **71**(4): 1017-1046.

Strathmann, M. F. (1987). Reproduction and development of marine invertebrates of the northern Pacific coast : data and methods for the study of eggs, embryos, and larvae. Seattle, University of Washington Press.

Stuart, T., A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert and R. Satija (2019). "Comprehensive Integration of Single-Cell Data." Cell **177**(7): 1888-1902 e1821.

Sun, Y., I. M. Dykes, X. Liang, S. R. Eng, S. M. Evans and E. E. Turner (2008). "A central role for Islet1 in sensory neuron development linking sensory and spinal gene regulatory programs." Nat Neurosci **11**(11): 1283-1293.

Sun, Z. and C. A. Ettensohn (2014). "Signal-dependent regulation of the sea urchin skeletogenic gene regulatory network." Gene Expr Patterns **16**(2): 93-103.

Sutherby, J., J. L. Giardini, J. Nguyen, G. Wessel, M. Leguia and A. Heyland (2013). "Histamine is a modulator of metamorphic competence in Strongylocentrotus purpuratus (Echinodermata: Echinoidea) (vol 30, pg 14, 2012)." Bmc Developmental Biology **13**.

Takacs, C. M., G. Amore, P. Oliveri, A. J. Poustka, D. Wang, R. D. Burke and K. J. Peterson (2004). "Expression of an NK2 homeodomain gene in the apical ectoderm defines a new territory in the early sea urchin embryo." Dev Biol **269**(1): 152-164.

Takei, K., V. I. Slepnev, V. Haucke and P. De Camilli (1999). "Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis." Nat Cell Biol **1**(1): 33-39.

Tan, D. X., L. C. Manchester, E. Esteban-Zubero, Z. Zhou and R. J. Reiter (2015). "Melatonin as a Potent and Inducible Endogenous Antioxidant: Synthesis and Metabolism." Molecules **20**(10): 18886-18906.

Tang F, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B Tuch, A. Siddiqui, K. Lao & M A. Surani. (2009). " mRNA-Seq whole-transcriptome analysis of a single cell". Nature Methods (6), 377–382

Tatsumi, T., N. Keira, K. Akashi, M. Kobara, S. Matoba, J. Shiraishi, S. Yamanaka, A. Mano, M. Takeda, S. Nishikawa, J. Asayama, H. Fliss and M. Nakagawa (2004). "Nitric oxide-cGMP pathway is involved in endotoxin-induced contractile dysfunction in rat hearts." J Appl Physiol (1985) **96**(3): 853-860.

Teitelman, G., S. Alpert, J. M. Polak, A. Martinez and D. Hanahan (1993). "Precursor cells of mouse endocrine pancreas coexpress insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide." Development **118**(4): 1031-1039.

Teng, F. Y., Y. Wang and B. L. Tang (2001). "The syntaxins." Genome Biol **2**(11): REVIEWS3012.

Thor, S., J. Ericson, T. Brannstrom and T. Edlund (1991). "The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat." Neuron **7**(6): 881-889.

Tian, J. H., Z. X. Wu, M. Unzicker, L. Lu, Q. Cai, C. Li, C. Schirra, U. Matti, D. Stevens, C. Deng, J. Rettig and Z. H. Sheng (2005). "The role of Snapin in neurosecretion: snapin knock-out mice exhibit impaired calcium-dependent exocytosis of large dense-core vesicles in chromaffin cells." J Neurosci **25**(45): 10546-10555.

Titze, B. and C. Genoud (2016). "Volume scanning electron microscopy for imaging biological ultrastructure." Biol Cell **108**(11): 307-323.

Tosches, M. A. and D. Arendt (2013). "The bilaterian forebrain: an evolutionary chimaera." Curr Opin Neurobiol **23**(6): 1080-1089.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold and L. Pachter (2010). "Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation." Nat Biotechnol **28**(5): 511-515.

Tritschler, S., F. J. Theis, H. Lickert and A. Bottcher (2017). "Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas." Mol Metab **6**(9): 974-990.

Tsironis, I., P. Paganos, G. Gouvi, P. Tsimpos, A. Stamopoulou, M. I. Arnone and C. N. Flytzanis (2021). "Coup-TF: A maternal factor essential for differentiation along the embryonic axes in the sea urchin Paracentrotus lividus." Dev Biol.

Tu, Q., C. T. Brown, E. H. Davidson and P. Oliveri (2006). "Sea urchin Forkhead gene family: phylogeny and embryonic expression." Dev Biol **300**(1): 49-62.

Tu, Q., R. A. Cameron and E. H. Davidson (2014). "Quantitative developmental transcriptomes of the sea urchin Strongylocentrotus purpuratus." Dev Biol **385**(2): 160-167.

Ushkaryov, Y. A., A. G. Petrenko, M. Geppert and T. C. Sudhof (1992). "Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor and laminin." Science **257**(5066): 50-56.

Valencia, J. E., R. Feuda, D. O. Mellott, R. D. Burke and I. S. Peter (2019). "Ciliary photoreceptors in sea urchin larvae indicate pan-deuterostome cell type conservation." bioRxiv: 683318.

Vult von Steyern, F., V. Martinov, I. Rabben, A. Nja, O. de Lapeyriere and T. Lomo (1999). "The homeodomain transcription factors Islet 1 and HB9 are expressed in adult alpha and gamma motoneurons identified by selective retrograde tracing." Eur J Neurosci **11**(6): 2093-2102.

Wagner, D. E., C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason and A. M. Klein (2018). "Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo." Science **360**(6392): 981-987.

Wang, Q., E. Mergia, D. Koesling and T. Mittmann (2017). "Nitric oxide/cGMP signaling via guanylyl cyclase isoform 1 modulates glutamate and GABA release in somatosensory cortex of mice." Neuroscience **360**: 180-189.

Wang, Y. J., J. Schug, K. J. Won, C. Liu, A. Naji, D. Avrahami, M. L. Golson and K. H. Kaestner (2016). "Single-Cell Transcriptomics of the Human Endocrine Pancreas." Diabetes **65**(10): 3028-3038.

Wei, Z., L. M. Angerer and R. C. Angerer (2016). "Neurogenic gene regulatory pathways in the sea urchin embryo." Development **143**(2): 298-305.

Wei, Z., R. C. Angerer and L. M. Angerer (2011). "Direct development of neurons within foregut endoderm of sea urchin embryos." Proc Natl Acad Sci U S A **108**(22): 9143-9147.

Wei, Z., R. Range, R. Angerer and L. Angerer (2012). "Axial patterning interactions in the sea urchin embryo: suppression of nodal by Wnt1 signaling." Development **139**(9): 1662-1669.

Wei, Z., J. Yaguchi, S. Yaguchi, R. C. Angerer and L. M. Angerer (2009). "The sea urchin animal pole domain is a Six3-dependent neurogenic patterning center." Development **136**(7): 1179-1189.
Wendik, B., E. Maier and D. Meyer (2004). "Zebrafish mnx genes in endocrine and exocrine pancreas formation." Dev Biol **268**(2): 372-383.

Wessel, G. M., M. Kiyomoto, T. L. Shen and M. Yajima (2020). "Genetic manipulation of the pigment pathway in a sea urchin reveals distinct lineage commitment prior to metamorphosis in the bilateral to radial body plan transition." Sci Rep **10**(1): 1973.

Wessel, G. M. and D. R. McClay (1987). "Gastrulation in the sea urchin embryo requires the deposition of crosslinked collagen within the extracellular matrix." Dev Biol **121**(1): 149-165.

Wikramanayake, A. H., L. Huang and W. H. Klein (1998). "beta-Catenin is essential for patterning the maternally specified animal-vegetal axis in the sea urchin embryo." Proc Natl Acad Sci U S A **95**(16): 9343-9348.

Wilfinger, A., V. Arkhipova and D. Meyer (2013). "Cell type and tissue specific function of islet genes in zebrafish pancreas development." Dev Biol **378**(1): 25-37.

Wilson, M. E., D. Scheel and M. S. German (2003). "Gene expression cascades in pancreatic development." Mech Dev **120**(1): 65-80.

Wollny, D., S. Zhao, I. Everlien, X. Lun, J. Brunken, D. Brune, F. Ziebell, I. Tabansky, W. Weichert, A. Marciniak-Czochra and A. Martin-Villalba (2016). "Single-Cell Analysis Uncovers Clonal Acinar Cell Heterogeneity in the Adult Pancreas." Dev Cell **39**(3): 289-301.

Wood, N. J., T. Mattiello, M. L. Rowe, L. Ward, M. Perillo, M. I. Arnone, M. R. Elphick and P. Oliveri (2018). "Neuropeptidergic Systems in Pluteus Larvae of the Sea Urchin Strongylocentrotus purpuratus: Neurochemical Complexity in a "Simple" Nervous System." Front Endocrinol (Lausanne) **9**: 628.

Ximerakis, M., S. L. Lipnick, B. T. Innes, S. K. Simmons, X. Adiconis, D. Dionne, B. A. Mayweather, L. Nguyen, Z. Niziolek, C. Ozek, V. L. Butty, R. Isserlin, S. M. Buchanan, S. S. Levine, A. Regev, G. D. Bader, J. Z. Levin and L. L. Rubin (2019). "Single-cell transcriptomic profiling of the aging mouse brain." Nat Neurosci **22**(10): 1696-1708.

Xu, F., F. Marlétaz, D. Gavriouchkina, X. Liu, T. Sauka-Spengler, G. Zhang, P.W.H. Holland. "Evidence from oyster suggests an ancient role for Pdx in regulating insulin gene expression in animals". Nat Commun **12,** 3117 (2021).

Yaguchi, J., N. Takeda, K. Inaba and S. Yaguchi (2016). "Cooperative Wnt-Nodal Signals Regulate the Patterning of Anterior Neuroectoderm." PLoS Genet **12**(4): e1006001.

Yaguchi, J. and S. Yaguchi (2019). "Evolution of nitric oxide regulation of gut function." Proc Natl Acad Sci U S A **116**(12): 5607-5612.

Yaguchi, J., A. Yamazaki and S. Yaguchi (2018). "Meis transcription factor maintains the neurogenic ectoderm and regulates the anterior-posterior patterning in embryos of a sea urchin, Hemicentrotus pulcherrimus." Dev Biol **444**(1): 1-8.

Yaguchi, S. and H. Katow (2003). "Expression of tryptophan 5-hydroxylase gene during sea urchin neurogenesis and role of serotonergic nervous system in larval behavior." J Comp Neurol **466**(2): 219-229.

Yaguchi, S., J. Yaguchi, H. Suzuki, S. Kinjo, M. Kiyomoto, K. Ikeo and T. Yamamoto (2020). "Establishment of homozygous knock-out sea urchins." Curr Biol **30**(10): R427-R429.

Yaguchi, S., J. Yaguchi and H. Tanaka (2017). "Troponin-I is present as an essential component of muscles in echinoderm larvae." Scientific reports **7**: 43563-43563.

Yaguchi, S., J. Yaguchi, Z. Wei, Y. Jin, L. M. Angerer and K. Inaba (2011). "Fez function is required to maintain the size of the animal plate in the sea urchin embryo." Development **138**(19): 4233-4243.

Yaguchi, S., J. Yaguchi, Z. Wei, K. Shiba, L. M. Angerer and K. Inaba (2010). "ankAT-1 is a novel gene mediating the apical tuft formation in the sea urchin embryo." Dev Biol **348**(1): 67-75.

Yajima, M., E. A. Gustafson, J. L. Song and G. M. Wessel (2014). "Piwi regulates Vasa accumulation during embryogenesis in the sea urchin." Dev Dyn **243**(3): 451-458.

Yajima, M. and G. M. Wessel (2011). "Small micromeres contribute to the germline in the sea urchin." Development **138**(2): 237-243.

Yamamoto, K. and P. Vernier (2011). "The evolution of dopamine systems in chordates." Front Neuroanat **5**: 21.

Yoshioka, Y., A. Yamamuro and S. Maeda (2006). "Nitric oxide/cGMP signaling pathway protects RAW264 cells against nitric oxide-induced apoptosis by inhibiting

the activation of p38 mitogen-activated protein kinase." J Pharmacol Sci **101**(2): 126-134.

Youson, J. H., A. A. Al-Mahrouki, Y. Amemiya, L. C. Graham, C. J. Montpetit and D. M. Irwin (2006). "The fish endocrine pancreas: review, new data, and future research directions in ontogeny and phylogeny." Gen Comp Endocrinol **148**(2): 105-115.

Yu, Z., J. Liao, Y. Chen, C. Zou, H. Zhang, J. Cheng, D. Liu, T. Li, Q. Zhang, J. Li, X. Yang, Y. Ye, Z. Huang, X. Long, R. Yang and Z. Mo (2019). "Single-Cell Transcriptomic Map of the Human and Mouse Bladders." J Am Soc Nephrol **30**(11): 2159-2176.

Yuh, C. H., X. Li, E. H. Davidson and W. H. Klein (2001). "Correct Expression of spec2a in the sea urchin embryo requires both Otx and other cis-regulatory elements." Dev Biol **232**(2): 424-438.

Yui, R. and T. Fujita (1986). "Immunocytochemical studies on the pancreatic islets of the ratfish Chimaera monstrosa." Arch Histol Jpn **49**(3): 369-377.

Zaninetti, M., E. Tribollet, D. Bertrand and M. Raggenbass (1999). "Presence of functional neuronal nicotinic acetylcholine receptors in brainstem motoneurons of the rat." Eur J Neurosci **11**(8): 2737-2748.

Zaret, K. S. and M. Grompe (2008). "Generation and regeneration of cells of the liver and pancreas." Science **322**(5907): 1490-1494.

Zecchin, E., A. Mavropoulos, N. Devos, A. Filippi, N. Tiso, D. Meyer, B. Peers, M. Bortolussi and F. Argenton (2004). "Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates." Dev Biol **268**(1): 174-184.

Zhao, D., Y. Yu, Y. Shen, Q. Liu, Z. Zhao, R. Sharma and R. J. Reiter (2019). "Melatonin Synthesis and Function: Evolutionary History in Animals and Plants." Front Endocrinol (Lausanne) **10**: 249.

Zhao, J., S. Zhang, Y. Liu, X. He, M. Qu, G. Xu, H. Wang, M. Huang, J. Pan, Z. Liu, Z. Li, L. Liu and Z. Zhang (2020). "Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human." Cell Discov **6**: 22.

Zhu, Y., Q. Liu, Z. Zhou and Y. Ikeda (2017). "PDX1, Neurogenin-3, and MAFA: critical transcription regulators for beta cell development and regeneration." Stem Cell Res Ther **8**(1): 240.

Zinovyeva, M. V., A. I. Kuzmich, G. S. Monastyrskaya and E. D. Sverdlov (2016). "[the Role of the Foxa Subfamily Factors in the Embryonic Development and Carcinogenesis of the Pancreas.]." Mol Gen Mikrobiol Virusol **34**(3): 98-103.